SESSION

LEARNING AND TEACHING METHODS, E-LEARNING + EDUCATIONAL TOOLS, AND RELATED ISSUES

Chair(s)

TBA

Students' Perception towards Soft CLIL in the Basque Secondary Schools

Chiharu Nakanishi[1] **and Hodaka Nakanishi**[2]
[1]Music Department, Kunitachi College of Music, Tachikawa, Tokyo, Japan
[2]Joint Program Center, Teikyo University, Itabashi-ku, Tokyo, Japan

Abstract - *The Basque Autonomous Community, which is a bilingual community of Basque and Spanish, has adopted CLIL as a method to teach English as a foreign language with on-line materials as well as written materials. The object of this study is to investigate how the students think about studying subjects in English, in the form of Soft CLIL (language-driven) in the Basque-medium school. The participants (n = 127) are 1st year secondary students. For them it is the 8th year to study English and 5th month to study subjects in English, in Soft CLIL. The results show that attitudes towards studying subjects in English are positive among overall students, whereas the majority of them want to learn subjects in Basque. The students who want to study subjects in English perceive the effect of Soft CLIL more than the others. The results obtained seem to confirm that Soft CLIL plays a positive role as a pre-stage to advance to Hard CLIL (Content-driven).*

Keywords: The Basque Autonomous Community, Soft CLIL, perception, attitudes

1 Introduction

In Europe, children are learning subjects such as history, geography and science in their second or third language in primary and secondary school. This has been partly due to a commitment of the European Union (EU) to a multilingual Europe. In 1995 EU proposed that every European citizen should master two other languages in addition to their mother tongue as one of the objectives of the language policy. Following this proposal, EU leaders at the March 2002 Barcelona European Council declared that every child in the EU should be taught at least two foreign languages from an early age. These requirements have given "CLIL" a great importance and have resulted in its establishment in many schools all around Europe. CLIL is an acronym of Content and Language Integrated Learning. Coyle et al. define "Content and Language Integrated Learning (CLIL) is a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language. That is, in the teaching and learning process, there is a focus not only on content, and not only on language. Each is interwoven, even if the emphasis is greater on one or the other in a given time (p.1) [1]."

As CLIL is an innovative approach, well-devised materials are to be developed. In order to meet the needs for teachers who want to implement CLIL in their classroom, several materials are provided on-line such as E-CLIL by European Resource Centre and EKI project in Basque whose materials are used in the schools we observed. EKI materials include music and video on Web which teachers use through internet in their classes as well as written textbooks [2].

In CLIL there are many types of programs and modules depending on the situation, context, and aim. According to Ball, there is a basic division of the CLIL approach into two camps: one in which the teaching and learning is focused primarily on language, and the other in which the teaching and learning is focused primarily on the subject content [3]. The former is referred to as language-driven, and is called "Soft CLIL" or "Light CLIL". The latter is called content-driven, and is often called "Hard CLIL" or "Heavy CLIL". Ball explains that the content-driven approach means that the subject content is given primary focus and subject concepts and skills as its learning objective, whereas a language-driven approach has as its basic objective language learning [3]. Students need gradual steps to move from Soft to Hard CLIL according to their progress of the target language.

2 CLIL in the Basque Autonomous Community

The Basque Autonomous Community (BAC), which is a bilingual community of Basque and Spanish, has adopted CLIL as a method to teach English. In the already crowded curriculum of Basque, Spanish and other subjects, integrating content and English is the most effective way to promote multilingualism. Under "the Eleanitz-project" which means multilingual in Basque, the BAC intends to educate and prepare students as future citizens of a multilingual and multicultural Europe [4]. "The Eleanitz-English" is part of the Eleanitz-project which integrates language policy with curriculum, materials and training for the teaching and learning of English for students aged 4-16. The Eleanitz-English project at "Ikastolas" ("Ikastolas" means schools in Basque), starts with an early introduction

of English at the age of 4 and lasts until the age of 16. During 4-years compulsory secondary school education, the students, who have studied English for 8 years already, experience to study subjects in CLIL. In the first two years, Year 1 and 2, they study subjects such as Music, Natural Science, and Technology in English classes of the standard secondary school curriculum from English teachers in the form of Soft CLIL. It is called "Subject Project". They have four English classes per week during "Subject Project". During the last two years, Year 3 and 4, they have two English classes and three Social Science classes in English per week. This project is called "SSLIC – Social Sciences and Language Integrated Curriculum" and they study under subject teachers in the form of Hard CLIL [4][5].

The educational system in the BAC has enhanced bilingual education and has taken important steps toward the development of multilingual education. During the Francisco Franco's dictatorship (in office: 1939 - 1975), the Spanish language was declared Spain's only official language and the public use of other languages was banned. Since the Status of Autonomy for the Basque Country was promulgated in 1979, one of the priorities of the Basque Government has been doing is the revitalization of the Basque language. When the Basic Law on the Standardization of Basque was passed in 1982, three linguistic school models were established to ensure that every student had the possibility to learn in Spanish and/or Basque [6][7].

Model A: Spanish medium schools

Model B: Spanish - Basque bilingual schools

Model D: Basque medium schools

Among the three linguistic school models, Model D (Basque medium schools) is the most dominant in all level of schools from kindergarten to university [5]. Some studies show that Model D can educate more balanced bilinguals of Basque and Spanish [8][9]. Lasagabster shows that the students enrolled in Model D scored higher points in Basque and English than those in other Models [10]. Regardless of medium of the instructive languages, all Models include English in their curriculum. It is based on the strong support from Government of Basque which promotes multilingual education. Cenoz states that parents are also strong supporters of multilingual education [6]. They are very interested in improving their children's level of proficiency in English [6]. Cenoz says that the development of multilingual programs in the BAC is related to the need to acquire English as a language of international communication and the idea that the positive experience of bilingual education can be extended to trilingual education [6].

While the BAC has been taking a step forward from a bilingual to multilingual community, CLIL has been implemented there. It is important to know the perception of students on CLIL. In the context of Basque and Spanish bilingual society, are the students willing to study subjects by the medium of an additional language, English? Lasagabaster examines the university students' attitudes toward English in the BAC. He finds that though they do not prefer to be taught in English, they think that English is worth learning and learning English enriches their cultural knowledge [11]. Lasagabster also examines secondary school students' attitudes in CLIL towards trilingualism [12]. He reveals that the CLIL can boost positive attitudes towards trilingualism at school, which leads to the multilingual educational system in Europe [12]. In the study of Lasagabster and Sierra, the CLIL students hold more positive attitudes towards English than those of EFL [13]. In both studies of Lasagabaster and Lasagabster & Sierra, the participants were 3rd and 4th year students of secondary school [12] [13]. The students in Lasagabaster's study studied subjects such as Geography, Arts or Social Science in English, that is, Hard CLIL [12]. In Europe, there are several studies focusing on perception and attitudes towards CLIL such as Dalton-Puffer et al. [14] and Yang & Gosling [15]. These studies tell the attitude towards CLIL is very important for the effective implementation of CLIL. However, the attitudes of Soft CLIL students in the BAC have not yet been investigated. It is necessary to find out how the Soft CLIL students think about their Soft CLIL in order to proceed to the step of Hard CLIL in a few years later.

The object of the present study is to investigate the perception of the secondary school students in the Basque-medium schools in Soft CLIL. The participants (n = 127) are 1st year students of five different schools in the BAC. They have studied English under the Eleanitz project at Ikastolas for 8 years. After entering secondary school, they started to study subjects in English, in Soft CLIL. The Soft CLIL classes are considered to be a pre-stage for taking Hard CLIL in Year 3 and 4.

3 The Present Study

3.1 Research Question

How do the 1st year students at secondary school think about studying subjects in English, Soft CLIL?

3.2 Method

3.2.1 Participants

A total of 127 students at 1st year of secondary school, 12-13 years-old, participated in this study. They were at five different Ikasotolas. The names of the Ikastolas and the number of students of the observed classes are as follows: Aranzadi (n = 22), Zurriola (n = 30), Axular Lizeoa (n = 26), Santo Tomas Lizeoa (n = 27), and San Benito (n = 22). All of these Ikastolas took Model D. The students started English at the age of 4 and it was the 8th year for their learning English under the Eleanitz-English project. They were in the period of "Subject Project" (Year 1 and 2). They studied the subjects in English class through the varying contents of different subjects such as Music, Natural Science, and Technology in Soft CLIL.

3.2.2 Materials

Soft CLIL materials of English in the Eleanitz project are prepared by the EKI project. The materials were content-led and employed the same type of conceptual sequencing and a part of the materials were e-materials and extra resources were supplied on the Web page.

3.2.3 Classes

At the time of the observation, it was the 5th month for the students after they started studying in Soft CLIL. The classes were taught by Basque native English teachers who had CLIL teacher training. Most of the time, the classes were managed in English. The teachers emphasized the students to interact with self, content-material, other students and teachers. The Web materials were used to show the concept of natural science and music. They were well devised and successful to attract the interest of students.

The participants would study Social Sciences and Language Integrated Curriculum, SSLIC, in their Year 3 and 4 under the subject teachers. It means that Year 1 and 2 in Soft CLIL, are in the pre-stage of Hard CLIL.

3.3 Instruments

The students were asked to fill out a questionnaire during class time of English in about 10 minutes. At the beginning of the questionnaire, they were asked their language circumstances i.e. languages at home, among friends, and language in which they wanted to study subjects. The questionnaire consisted of a question of general preference of English and four questions of effect of Soft CLIL with a 5-point Likert scale (Table 1). "1" is strongly disagree, and "5" is strongly agree. While students were filling out the questionnaire, the English teachers and the authors of this paper helped them.

Table 1: Questionnaire

Items		Questions
General preference		I like studying subjects in English very much.
Effect of Soft-CLIL	English proficiency	Studying subjects in English improves my English proficiency very much.
	Understanding content	Studying subjects in English improves my understanding of the content very much.
	Thinking skills	Studying subjects in English improves my thinking skills very much.
	World	Studying subjects in English opens my door to the world.

4 Results

All answers to the questions were analyzed from the viewpoint of their language circumstances, general preference of English, and effect of Soft CLIL.

4.1 How do the Year 1 students think about Soft CLIL?

4.1.1 Language circumstances

Language circumstances of the students are shown in Table 2. Among 127 students, the highest ratio of language at home is Basque only (48.0%) following both Basque and Spanish (28.3%), and Spanish only (22.8%). Among friends, the tendency is the same. Basque is the most spoken (44.1%), followed by both Basque and Spanish (39.4%), and Spanish (15.7%).

4.1.2 Preferred language to study subjects

The preferences of language for studying subjects are shown in Table 3. The number of the students who want to study subjects in Basque is the majority (77.2%), followed by Spanish (15.7%). The students who want to study subjects in English is only 14.2%

4.1.3 Effect of CLIL

The students' perceptions towards effect of CLIL are shown in Table 4. Although the differences of average points among the items except "World" are not significant, the average points of "World" are significantly high comparing to the other points of items ($p < .05$). As the average points of items are higher than 3 (i.e. middle), the majority of the students enrolled in Soft CLIL show positive attitudes towards studying subjects in English.

4.1.4 Differences between the students who prefer to study subjects in English and either in Basque or Spanish

The average points of each item by the students who want to study subjects in English and either in Basque or Spanish, are shown in Table 5. For the calculation of *p*-values, we use one-sided tests because the students who want to study in English are supposed to make higher evaluation on the effect of Soft CLIL. Based on the *F*-test, all questions but "World" are tested supposing the same distribution and "World" is tested supposing different distribution.

The *t*-test shows statistically significant differences between the students who want to study subjects in English and those who want to study in Basque or Spanish in the item of general preference towards studying subjects in English, and in the 3 items on the effect of Soft CLIL. Those students who want to study subjects in English think that

Soft CLIL improves their English proficiency, their understanding of content, and it opens their door to the world. However there is no significant difference in the item of thinking skills. The students who want to study subjects in English show more positive attitudes towards Soft CLIL than those in Basque or Spanish. However the average points obtained from two groups are higher than 4 except the item of general preference by those who want in Basque or Spanish. The majority of the students enrolled in Soft CLIL shows positive attitudes towards studying subjects in English.

4.1.5 Differences among the native languages towards Soft CLIL

The points obtained of each item by the native languages are shown in Table 6. The perceptions of Soft CLIL are significantly indifferent depending on either monolingual of Basque or Spanish, or bilingual.

Table 2: Language at home and among friends (n = 127)

	Basque only	Spanish only	Basque & Spanish	Others
Language at home	48.0% (n = 61)	22.8% (n = 29)	28.3% (n = 36)	0.8% (n = 1)
Language among friends	44.1% (n = 56)	15.7% (n = 20)	39.4% (n = 50)	0.8% (n = 1)

Table 3: Preferred language to study subjects (Multiple choice allowed) (n = 127)

	Basque	Spanish	English
Language in which they want to study subjects	77.2% (n = 98)	15.7% (n = 20)	14.2% (n = 18)

Table 4: Students' perception towards effect of CLIL (n = 127)

Items		Average Points (SD)
General preference of English		3.92 (.86)
Effect of Soft CLIL	English proficiency	4.19 (.70)
	Understanding content	4.15 (.83)
	Thinking skills	4.09 (.73)
	World	4.52 (.87)

Table 5: Differences between the students who want to study subjects in English and those who want to study subjects either in Basque or Spanish (n = 127)

Questions		Want to study in English (n = 18)	Want to study in Basque or Spanish (n = 109)	t - value
General preference		4.33	3.85	<.05
Effect of Soft CLIL	English proficiency	4.39	4.15	<.1
	Understanding content	4.44	4.11	<.1
	Thinking skills	4.22	4.07	
	World	4.78	4.48	<.05

Table 6: Differences among the native languages towards Soft CLIL (n = 127)

		Monolingual		Total of Monolingual	Bilingual	Total
		Basque	Spanish			
General preference		3.89	3.80	3.88	3.95	3.92
Effect of Soft CLIL	English proficiency	4.28	3.80	4.18	4.19	4.19
	Understanding content	4.14	3.78	4.07	4.19	4.15
	Thinking skills	4.22	3.90	4.16	4.05	4.09
	World	4.47	4.30	4.44	4.56	4.52

5 Conclusions

The present study showed that the 1st year students at secondary schools in the BAC did not prefer to study subjects in English but preferred to study in Basque, their mother tongue. It should be easy for them to study subjects in their native language. Although only 14.2% of the students preferred to study subjects in English, the majority of the

students perceived the effect of Soft CLIL positively. They perceived Soft CLIL was effective to improve their English proficiency, understanding the content, and it opened the door to the world.

The gap between perception of Soft CLIL and preference of language might be because a sense of insecurity for CLIL still remains in students as Soft CLIL has been introduced for only five months. One of the reasons of positive attitudes on the effect of Soft CLIL is that *"the Eleanitz-project and the Eleanitz-English"* in which the students participate are said to be successful. Elorza and Lindsay (2013) describe one of the common perspective in the *Ikastolas* is "Language is learnt as we use it as a tool for life in natural contexts of communication". Since this perspective was observed in their English classes, the students may feel that they can use English naturally and authentic way.

Among four questions of effect of Soft CLIL, the perception towards item of "World" was the highest. The students may value that studying English links to the outside of the classroom and world. The value of *"the Eleanitz-project"* and parents who support English and promote multilingualism (Cenoz, 2005), might mirror the students' positive attitudes towards learning English in CLIL. The results obtained seem to confirm that Soft CLIL plays a role as a pre-stage to advance to Hard CLIL in later stage of language curriculum.

There are some limits of the present study. The 5 questions are not enough to do factor-analysis. As a teacher, we know that the maximum time limit for students to answer a *questionnaire* in a class is about 10 minutes. The questionnaire which needs longer than 10 minutes, does bother students and is not welcomed most of the time.

For our future study, we'd like to investigate how the 1st Year students' perception towards Soft CLIL will change when they finish *"Subject Project"*, at Year 2, and how they think about Hard CLIL at Year 3 and 4. We'd like to see the transition or difference between Soft and Hard CLIL in their perception in a long-term research. Moreover we are also interested in the correlation between the achievements and perception towards CLIL.

6 Acknowledgements

This work was supported by Grants-in-Aid for Scientific Research, Grant Number 25350294, awarded by Japan Society for the Promotion of Science. We show our special thanks to Phil Ball who supported all our research in the BAC. We are grateful for Noboru Wakayama, our colleague and the English teachers and students in the BAC who allowed us to observe the class and answered our questionnaire.

7 References

[1] D. Coyle, P. Hood & D. Marsh. "Content and Language Integrated Learning". Cambridge: Cambridge University Press, 2010.

[2] Ikaselkar argitaletxea. "EKI Proiektua", 2013. <http://www.ekigunea.com/>, Viewed on 15 March 2014.

[3] P. Ball. "How do you know if you' practising CLIL?" 2008. <http://www.onestopenglish.com/clil/methodology/articles/article-how-do-you-know-if-youre-practising-clil/500614.article>, Viewed on 15 August 2013.

[4] www.eleanitz.org. "Eleanitz Project", 2013. <http://www.eleanitz.org/public/Eleanitz_Project>, Viewed on 15 March 2014 .

[5] I. Elorza & D. Lindsay. "The Basque Education System & The Ikastola Network's Plurlingual Plan", Ikastolen Etxea, Donostia. 17th June, 2013, At University of the Basque Country, Symposium, Summer Courses 2013, 2013.

[6] J. Cenoz. "English in bilingual programs in the Basque Country", International Journal of the Sociology of Language, 171, pp. 41–56, 2005.

[7] R. Y. Zarobe & D. Lasagabaster. "CLIL in a bilingual community: The Basque Autonomous Region". In: CLIL in Spain: Implementation, Results and Teacher Training (eds. Lasagabaster, D. & Ruiz de Zarobe, Y.), 2010.

[8] J. J. Gabiña, R. Gorostidi, R. Iruretagoiena, I. Olaziregi, & J. Sierra. "EIFE:Influence of Factors on the Learning of Basque". Vitoria-Gasteiz: Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia, 1986.

[9] J. Sierra & I. Olaziregi. "EIFE 2: Influence of Factors on the Learning of Basque". Vitoria-Gasteiz: Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia, 1989.

[10] D. Lasagabaster. "Basque, Spanish and English: Three languages in contact in the Basque educational system", In Proceedings of the First International Symposium on Bilingualism, pp. 1078-1085, Vigo : University of Vigo, 2013.

[11] D. Lasagabaster. "Attitudes towards English in the Basque Autonomous Community", World Englishes, Vol. 23, pp. 211-224, 2004.

[12] D. Lasagabaster. "The implementation of CLIL and attitudes towards trilingualism". ITL, International Journal of Applied Linguistics, 157, pp. 23-43, 2009.

[13] D. Lasagabaster & J. M. Sierra. "Language Attitudes in CLIL and Traditional EFL classes". International CLIL Research Journal, Vol. 1 (2), *pp.* 4-17, 2009.

<http://www.icrj.eu/12/article1.html>, Viewed on 15 March 2014.

[14] C. Dalton-Puffer, J. Hüttner, V. Schindelegger & U. Smit. "Technology-Geeks Speak Out: What Students Think About Vocational CLIL", International CLIL Research Journal, Vol. 1 (2), pp.18-25, 2009. <http://www.icrj.eu/12/article2.html>, Viewed on 15 March 2014.

[15] W. Yang & M. Gosling. "National Appraisal and Stakeholder Perceptions of a Tertiary CLIL Programme in Taiwan", International CLIL Research Journal, Vol. 2 (1), pp.67-81, 2013. <http://www.icrj.eu/21/article6.html>, Viewed on 15 March 2014.

Exploring the Requirements and Infrastructure to Develop Online Degree Programs

Ingrid A. Buckley[1] **and Hira Narang**[1]
[1] Computer Science Department, Tuskegee University, Tuskegee, Alabama, USA

Abstract - *This paper presents the fundamental aspects that must be considered to develop online programs. The pros and cons associated with online education are discussed. The resources and requirements (infrastructure, software, tools, and training) needed to develop online degree programs are discussed. This paper provides a survey of tools and software programs available in both the private and public domains. We discuss the fundamental requirements, and the overall effort of implementing online degrees in terms of technology, infrastructure, online course development and faculty training.*

Keywords: Online Degree, Course Management Tools, Course Management Software, Online Education, Online Infrastructure

1 Introduction

Online education has seen dramatic growth over the last few decades. Online education has significantly increased in popularity among students of all ages. Online learning allows universities to increase their enrollment by creating distance learning programs. Distance learning programs provide a flexible and convenient way for students to further their tertiary education. Online courses can be accessed by students at any time without decreasing the quality of education that is achieved in on-site degree programs or online courses. Today, online learning has become a mandatory addition to many universities degree programs because, it affords students from varying backgrounds and circumstances the opportunity to further their education [23]. Many universities [3] have leveraged the benefits of online education and students are responding positively to this change and reaping the many benefits.

The paper is organized in the following manner. Section 2 discusses some pedagogical pros and cons of online education. Section 3 presents infrastructure needs, educational tools, commercial and open source software for managing online programs, along with various institutions and companies that develop online course content. Section 4 provides a cost and benefit analysis. Section 5 presents the groundwork to implement an online degree program. The paper concludes with section 6.

2 Advantages and Disadvantages of Online Education

Online Education was not originally popular; however, the internet and the ubiquity of technology have impacted online learning positively. Online Education is responsible for increased student enrollment and access to higher education from anywhere across the globe. There are some fundamental reasons that support the viability of online education. First, millennials typically choose non-traditional education to start and advance their careers while completing and furthering their education. Second, the down-turn in the US economy forced people to upgrade/change their career through online educational programs.

With the increased popularity of online education, there are notable advantages and disadvantages that result when compared to the traditional education system. The advantages include convenience, reduced cost and flexibility of available technology [24]. In particular students are not restricted by their geographic location, they do not have commute and pay house expenses, and they can access school work from anywhere. The increased access also allows universities and international students to avoid visa and immigration problems. Online education also presents some disadvantages, namely inadequate social interaction with peers, technology cost and assessment and instructor issues. Online students may experience a lack of face-to-face dialog with instructors, high dependence on high speed internet, and may have attain technical trouble shooting skills to handle technical problems. Additionally, traditional instructors accustomed to handouts may have a difficulty adapting to the online systems and software.

Note that some of the disadvantages mentioned are also experienced in the traditional education system. Typically, unmotivated students who do not show up for class usually do not perform well. Exposing students to new technology and troubleshooting issues is an additional benefit, as it allows them to master the technology they have to work with, these skills will serve them well beyond the online classroom [24]. Faculty members who have rooted themselves in the sanctuary of campus based face to face education are the most vocal critics of online education.

3 Infrastructure and Software Needs for Online Education

We discuss and summarize some of the infrastructure and software needs that are required to develop an online degree program.

3.1 Online Infrastructure

All educational endeavors are systems, made up of various interconnected components. In distance education, understanding how the entire system of course development and delivery occurs, and how these systems link to services and other components are important aspects of ensuring effectiveness and quality. In order to create high quality videos, a controlled and well equipped environment is necessary to record online classes. The various communication mediums have to be clear and succinct without delay and distortion, because students rely heavily on excellent communication to carry out their studies remotely.

Additionally, SMART Boards [35] have been adopted by many universities since they combine the simplicity of a whiteboard with the power of a computer. The SMART Board interactive whiteboard allows for the delivery of dynamic lessons, write notes in digital ink and save your work – all with the simple touch of a finger. The SMART Board uses an integrated SMART projector which limits shadows and projects quality image, making visuals crisp and bright.

Servers [21] are required to store and retrieve recorded videos and any other course material. Power Supply system sources should always be available to power the server and its associated sub-systems (cameras, SMART Board and mikes) to ensure that the content is not lost.

3.2 Software and Tools

Online education relies heavily on special tools and technologies in order to create a successful and functional learning environment. There are a variety of tools on the market that provide the resources needed by instructors and students. Wimba [40] and Blackboard [7] are two widely used products for supporting online learning.

Wimba provides educators and students the ability to build relationships by combining interactive technologies with instructional best practices. Wimba contains many capabilities which include support of audio, video, application sharing, and content. Wimba has advanced features such as polling, whiteboarding, presenter on-the-fly, resizable chat areas and participant lists, usage analytics tools. Wimba uses Voice over IP to afford natural communication; it modulates tone of voice, captures body language and conveys the natural enthusiasm and interest of the professor.

Wimba affords flexibility in that, a wide variety of cameras used on computer systems or devices will be compatible with Wimba. This allows students to obtain the benefits of online instruction without the hassle of incompatibility issues. This reduces help desk support for students. Additionally, Wimba provides SSL encryption for classroom traffic thus reducing disruption from outside intruders whilst upholding the standards of learning.

Wimba can be easily integrated with Blackboard through Blackboard Collaborate. *Blackboard Collaborate* [8] manages pre-session planning and post-session recording management tools that support the entire instructional cycle. Blackboard collaborate allows students and professors to view classroom recordings on a desktop or easily convert for download or streaming to any mobile device to meet the needs of on-demand active learners.

Blackboard provides institutions the ability to develop and implement a learning management system that impacts every aspect of education. Blackboard allows professors to engage students in exciting new ways, reaching them on their terms and devices - and connecting more effectively, keeping students informed, involved, and collaborating together.

ANGEL Learning Management Suite (LMS) [4] is a tool for creating virtual learning environments for online learning. A learning management system (LMS) is a software application for the administration, documentation, tracking, and reporting of training programs. ANGEL is used to offer hybrid or blended (web-enhanced) classes. In addition to creating courses, and adding content, ANGEL also has features such as pattern recognition of online student activity, assessment tools that measure learning outcomes against class objectives.

Mentor Course Management System (CMS) [6] provides tools to manage work outside of class or teach fully online courses. The Mentor platform includes functionalities for course management, assessment, tests, grants management and faculty evaluations.

eFront [17] allows instructors to create of online learning communities while offering various opportunities for collaboration and interaction through an icon-based user interface. The platform includes tools for content creation, tests building, assignments management, reporting, internal messaging, forum, and surveys.

Sakai Collaboration Learning Environment (CLE) [33] is a free learning management, research and project collaboration, e-portfolio source. Sakai uses tools, for maintaining Learning Management System (LMS) learning management platforms such as presentation, profile, schedule, assignment, gradebook, assessment, syllabus, forms, glossary and reports.

ILIAS [22] is an open source Learning Management System (LMS) which offers features to design and run online-courses. These features include creating learning content, offer assessments and exercises, run surveys and support communication and cooperation among users. It contains modules like course management, groups, virtual classroom, chat, discussion and forums.

Online Learning And Training (OLAT) [28] is an open source Learning Management System (LMS) which is available in a variety of languages and provides diverse functionalities in web-based learning and training. OLAT includes standard modules for learning delivery and course management like content managing, forums, file discussions, quizzes, logs, podcast, assignment submissions and e-assessments.

3.3 Companies Offering Online course Development Support

Developing an online course requires a different way of planning, presenting and delivering course content to distant learners. Online course development requires new skills. Several institutions and companies [19] provide assistance with online course development**.** Higher Education institutions offering courses online to meet increasing demand often times, procure outside help to develop or design their digital curriculum. The American Distance Education Consortium (ADEC) [2] is a non-profit distance education consortium. The consortium was developed to encourage the creation and delivery of quality, economical distance education programs and services to diverse audiences. ADEC utilizes subject matter specialists and information resources to share knowledge and content with learners. ADEC offers assistance locally, regionally, nationally, and internationally. Their focus includes helping institutions with design for active and effective learning, develop and maintain the technological and human infrastructure [2].

Accel Media [1] provides training and services centered on online course development and assessment. They convert existing traditional courses into online courses, create instructional curriculum design, voice-over narration, art and graphics. Additionally, Accel Media provides technical support to users, host, manage and update online courses. *Mindflash* [25] uses existing training content to build interactive web-based courses in a timely fashion. They include features and tools for training, deployment, and progress management for online courses. *TestOut* [36] offers interactive and intuitive tools that support effective communication and delivery of technical courses such as networking, CISCO certification and security online.

Velsoft [38] provides a courseware that includes content and training material that allows institutions to create, customize and manage mobile friendly, non-technical online courses. *Skye Multimedia* [34] offers a virtual web based technology, called Reflect, that can be combined with classroom training to reinforce traditional teaching methods. *Reflect* [32] provides a broad training solution by taking learners through a cycle of instruction, demonstration, practice, review and feedback. *Digital Latitude* [16] offers custom E-Learning course development and instructional design for a variety of companies. Their product offers off-the-shelf, semi and custom courses, in addition to software that monitors the progress made by participants in each course.

Colloquy [13] offers online course development solutions to facilitate the financial and technical challenges institutions experience when taking their courses online. Colloquy also offers marketing and recruiting, curriculum design, development, learning management systems, and student success services. *The Learning House* [37] helps universities to create, manage and grow online degree programs and courses. They also provide the following services: market research, marketing and lead generation, admissions and enrollment management, student retention, curriculum development and management, faculty training, professional development, learning management systems and 24/7 technical support.

3.4 Open-Source Software for Online Course Development

There are several open source software available for online course development and for customizing courses using predesigned course templates. An existing course can be enhanced with the addition of embedded videos, links and tasks. These software are most suitable for descriptive or non-computer science courses. However, they can allow for the development of needed computer science courses within the available framework. The two most famous are: WebQuest [39] and HyperInquiry [20].

WebQuest [39] uses a bottom up approach to develop non-technical courses online. A WebQuest is **a**n inquiry-oriented lesson format in which most or all the information that learners work with comes from the internet. It consists of embedded videos, links, assessment, and information surrounding a topic. WebQuest has a variety of different course templates that can be used by instructors to quickly and effectively develop an online course.

HyperInquiry [20] uses an inductive approach to design non-technical courses online. HyperInquiry is based on a three phase model which encapsulates concept, interpretation of data and application of principles. This approach encourages students to understand, extrapolate, and apply information by coaching them to identify patterns for learning different topics.

Coursera [14] is an educational technology open source organization, which offers massive open online courses (MOOC) [31]. Coursera works with other universities to make some of their courses available online in engineering,

humanities, medicine, biology, social sciences, mathematics, business, computer science, and other areas. Each course includes short video lectures on different topics and assignments are submitted on a weekly basis. Web forums are provided for courses, and students can arrange face to face study meet-ups using meetup.com, or online meetups. Currently, Coursera has partnerships with 107 institutions and universities.

Course Builder [15] is new open-source software released by Google**.** Course Builder is a tool that universities can use to deliver free online courses. Course Builder is a joint partnership between Google, Harvard University, the Massachusetts Institute of Technology, and the University of California at Berkeley to offer free online courses. Google tested the software in summer 2013, by offering a course called Power Searching, which attracted 155,000 registered students. Course Builder is available for download, but its installation requires strong technical skills and a server to run it on.

iTunes U [5] is a new application software created by Apple, that gives instructors the capability of creating courses for students on an IPAD. The courses can be created using multiple formats such as audio videos, presentation, documents, and iBooks for ipads. It allows students to integrate with iBooks and other apps thus allowing them to keep abreast with courses. iTunes U app stores documents, notes, highlights, and bookmarks up to date across multiple devices.

Moodle [26] is an open-source alternative to Blackboard. Moodle is a free learning management system which provides functions for assignment submission, discussion forum, files download, grading, online quizzes and announcement.

OpenClass [29] is a free course-management system, which is completely Cloud based and is available under an open-source license. OpenClass provides a dynamic learning environment that helps educators bring social learning and experiences to students. OpenClass allows instructors to create courses, add content and manage course in one place. Students can access content via a computer or via mobile devices.

Canvas [11] is an open-source learning management system. It allows instructors to create assignments, share course information, and integrate assignments in the course calendar. It allows students to integrate their accounts with social media sites such as Twitter and Facebook. Individual Students and instructors who are not affiliated with a school can use the program for free.

4 The Foundation to Implement an Online Degree Program

Based on demand and the changing dynamics of universities globally, and the goal to attract more students, it may be worth while for smaller traditional universties to offer online degrees completely online at the undergraduate and graduate level. There are some significant foundational units that must be implemented before the development of an online degree program begins. The following guidelines show the fundamental units that are required to meet this objective:

4.1 Form Committees

It is important to establish administration, curriculum and course development, technology, finances and committees [27]. These committees will be responsible for:

- Identifying important participants, identify subject matter experts, and instructional designers.
- Leveraging the experiences and knowledge of individuals from a number of different disciplines and knowledge bases.
- Ensuring that individuals have unique and compatible skills sets.
- Establishing ground rules on how to share ideas and debate topics.

4.2 Collaboratively Develop Course Content

There are several models that can be used to accomplish this objective as outlined in [18], the process includes:

- Exchanging ideas and determining what content is needed.
- Developing assessments for the different course objectives.
- Selecting the appropriate resources for the course.

4.3 Develop Rubrics

A rubric can serve several beneficial purposes. A rubric can be used in a course as a self-evaluation tool to advise instructors how to revise an existing course. It can also be used for the design of a new course for the online environment by following the rubric as a road map [10]. Develop an online rubric:

- After the assignments have been developed collaboratively among the team, the rubric can be developed.
- The development of rubrics will standardize certain areas of the content that will be assessed which are connected to course objectives.

- The validity and reliability of the rubric must be evaluated.

4.4 Assessment

The primary function of the assessment is to measure student learning. There are two types of assessment, formative and summative. Formative happens while the instruction is in process, while summative is a final overall determination [30]. The assessment committee is tasked with the following responsibilities:

- Assuring that learning objectives are covered.
- Ensuring that the needed changes/ improvements to course assignments are done.
- Ensuring that all program outcomes are considered.
- Analyzing how the interactive learning opportunities were experienced by students.
- Determining if assignments adequately cover student learning outcomes.
- Determining what was missing from the course.

5 Analysis

One primary goal of universities that offer online degrees is to ensure quality and flexible education that is consistent with the current educational climate. More universities and community colleges are offering blended and online courses to meet the needs of the students they target. In order for institutions to remain competitive it may be beneficial to consider meeting the needs of diverse students. Online courses are geared towards student needs on a broader platform, this can increase student enrollment. The following aspects are essentials to develop an online degree program:

5.1 Initial Technology and Infrastructure Needs

There is an increasing range of user interfaces, physical devices and supporting infrastructure that facilitate online learning. It is important to meet with different information and communications technology (ICT) infrastructure providers [21] to get an estimate of the overall cost to develop an online program on a scale that is consistent with the requirements [9, 12]. Technology and infrastructure needs can become quite costly, however, there is the option of procuring the services of companies that offer the technology, server space, databases storage and online facilities that will support our online courses.

Most universities use some form of course management system, such as Blackboard for example. Integrating video lectures into the existing course management software used in your university is encouraged to reduce cost. Video lectures are useful as students can access them at any time. Generally, the infrastructure and technology needs in most universities may include a Cloud, some workstations and the appropriate software (operating system and applications) to develop and manage the online program.

5.2 Course Development and Faculty Training

Effective online course development, management and training will determine the success of online courses. The strength of online education is not only contingent on the medium it uses but in the way it is used. Similarly, online course development and training is provided by companies and other institutions at a cost. It is important to train faculty and staff members by sending them to various workshops, conferences, and professional seminars. Strong consideration must be given to training the faculty specifically for online teaching as well as the time allocated for training. Instructional design support and guidelines can help instructors get acclimated to this new form of teaching. The trained instructors can be utilized as mentors for new faculty members for future training or professional development efforts.

6 Conclusions

The rapid educational changes and learning needs of students require the enhanced benefits that online education offers. These benefits allow universities to expand and diversify their student body as well as their degree programs. However, the cost including infrastructure, faculty training and maintenance, must be thoroughly evaluated before implementing an online degree program. The benefits of online education include facilitating existing residential students, who may complete their degrees earlier because of the availability of certain online courses. Due to the success of universities who offer online degree programs a wide variety of effective tools, and course management software are now available free of charge. Additionally, universities should also consider the challenge online degrees have on instructors who are accustomed to teaching in the traditional class format. Instructors will have to learn the tools, course management software and find innovative approaches to convey knowledge through the online platform. In order to advance in pedagogy, and remain competitive it may be worthwhile for universities to at least explore the online direction.

7 References

[1] Accel Media. (2014). We Build Online Courses. Retrieved from: http://www.accelemedia.com/about.html.

[2] ADECO. (2013). "The American Distance Education Consortium". Retrieved from: http://www.adec.edu/.

[3] American School Search. (2014). "Colleges and Universities", Retrieved from: http://www.american-school-search.com/colleges/hbcu.

[4] ANGEL LMS. (2014). "ANGEL Learning Management Suite". Retrieved from: http://www.angellearning.com/products/lms/whats_new_74.html.

[5] Apple. (2014). "iTunes U". Retrieved from: http://www.apple.com/education/ipad/itunes-u/.

[6] AxiomEducation. (2014). "Course Management". Retrieved from: http://www.axiomeducation.com/manage-courses-curriculum-management/.

[7] Blackboard. (2014). "About BB". Retrieved from: http://www.blackboard.com/.

[8] Blackboard Collaborate. (2014). "About Blackboard Collaborate". Retrieved from: http://www.blackboard.com/platforms/collaborate/overview.aspx .

[9] Kerry Blinco, Jon Mason, Neil McLean and Scott Wilson. "Trends and Issues in E-learning Infrastructure Development", White paper, July 19, 2004. Retrieved from: http://www.jisc.ac.uk/uploaded_documents/Altilab04-infrastructureV2.pdf.

[10] California State University. (2009). "Rubric for Online Instruction". Retrieved from: http://www.csuchico.edu/roi/the_rubric.shtml.

[11] Canvas. (2014). "The University of Texas at Austin", Retrieved from: http://canvas.utexas.edu/.

[12] Center of Digital Education and Converge. (2013). "Smart Infrastructure", Special Report. Retrieved from: http://images.erepublic.com/documents/CDE13+SPQ1_V.PDF.

[13] Colloquy. (2014). "Solutions Overview". Retrieved from: http://colloquy360.com/.

[14] Coursera.(2013). "Our Mission". Retrieved from: *https://www.coursera.org/*.

[15] Course-Builder. (2014). "Welcome to Course Builder!" Retrieved from: https://code.google.com/p/course-builder.

[16] Digital Latitude. (2014). "Intelligent e-learning solutions". Retrieved from: http://www.digital-latitudes.com/.

[17] eFront.(2014). "5 Reasons to choose eFront". Retrieved from: http://www.efrontlearning.net/.

[18] Emily Hixon. "Team-based online course development: A case study of collaboration models", Online Journal of Distance Learning Administration, 2008. Retrieved from: http://www.westga.edu/~distance/ojdla/winter114/hixon114.html

[19] Hobson. (2014). "Higher Ed Institutions". Retrieved from: http://www.hobsons.com/education-professionals/higher-education/.

[20] HyperInquiry. (2014). "Instructional & Learning Strategies". Retrieved from: http://www.southalabama.edu/coe/bset/dempsey/isd613/summer04/unit/wq/hyperinquiry.html.

[21] IBM. (2013). "Buy BladeCenter HX5 Featured Models". Retrieved from: http://www-304.ibm.com/shop/americas/content/home/store_IBMPublicUSA/en_US/evp/hx5.html.

[22] ILIAS. (2014). "Using ILIAS". Retrieved from: http://www.ilias.de/docu/goto_docu_cat_580.html.

[23] Daphne Koller. "How online education can create a 'global classroom", CNN special report, June 21, 2013. Retrieved from: http://edition.cnn.com/2013/06/21/business/opinion-koller-education-petersburg-forum.

[24] Dhirendra Kumar. "Pros and Cons of Online Education", Manuscript, 2010, North Carolina State University, Raleigh, NC, USA.

[25] MindFlash. (2013). "Easy Online Training". Retrieved from: http://www.mindflash.com/.

[26] Moodle. (2014). "About Moodle". Retrieved from: http://docs.moodle.org/26/en/About_Moodle.

[27] Priscilla O. Okunji and Mary H. Hill, "Undergraduate Online Program Development, Implementation and Evaluation: A Pilot Study", Canadian Journal of Nursing Informatics(CJNI), 8, 3, September 23, 2013.

[28] OLAT.(2014). "Welcome to OLAT". Retrieved from: http://www.olat.org/.

[29] OpenClass. (2014). "Who Uses OpenClass?". Retrieved from: http://www.openclass.com/who .

[30] Joanna Palmer, Donna DiMatteo-Gibson, and Linda Phillips. "Building a Stronger Online Course/Program: Implementing a Six Step Collaborative Development Model", Sloan-C Conference 2013 International Conference on Online Learning, 2013. Retrieved from: http://sloanconsortium.org/conference/2013/aln/building-stronger-online-courseprogram-implementing-six-step-collaborative-devel.

[31] Laura Pappano. "The Year of the MOOC". The New York Times", November 2, 2012. Retrieved from: http://www.nytimes.com.

[32] REfLECT. (2014). "REfLECT classic". Retrieved from: http://www.ifl.ac.uk/cpd/reflect/reflect-classic.

[33] Sakai. (2014). "Explore the Sakai CLE". Retrieved from: http://www.sakaiproject.org/.

[34] Skye Multimedia. (2014). "iReflect". Retrieved from: http://www.skyemm.com/products.html ?product=ireflect#header_placeholder.

[35] SMART. (2013). "Smart Boards". Retrieved from: http://smarttech.com/smartboard.

[36] TestOut. (2014). "LabSim Educational Solutions". Retrieved from: http://www.testout.com/home/educator-resources/labsim-educational-solutions.

[37] The Learning House Inc. (2014). "what we do". Retrieved from: http://www.learninghouse.com/.

[38] Velsoft. (2014). "Velsoft". Retrieved from: http://www.velsoft.com.

[39] WebQuest. (2014). "What is a WebQuest?". Retrieved from: http://webquest.org/index.php.

[40] Wimba. (2014). "Wimba Classroom for Higher Education". Retrieved from: http://www.wimba.com/solutions/higher-education/wimba_classroom_for_higher_education.

Building a Frame-Based Content Generation Approach for APT Prevention Education

Yuan-Hsin Tung[1], Shian-Shyong Tseng[*2], Yuh-Jye Wang[2], Hwai-Ling Shan[1]

[1]*Telecommunication Lab., Chunghwa Telecom Co., Ltd., Taiwan,ROC*

yhdong@cht.com.tw, shanhl@cht.com.tw

[2]*Department of Applied Informatics and Multimedia, Asia University, ROC*

sstseng@asia.edu.tw, crazycavalier@gmail.com

***Abstract*—*In the recent years, there have been growing incidences of advanced persistent threats (APTs) that collect information and data on specific target of enterprise and then perform hacking by using the data. It results in a serious problem of causing enterprise heavy financial and political losses. As we know, the most efficient way to fight APTs is to educate the participants of the organizations with well APT prevention knowledge. Due to low motivation of traditional course contents, game-based learning utilizing interactive activities and interesting multimedia is a suitable solution for motivating end-users. However, it is usually cost consuming to generate numerous game contents of a full assortment to satisfy pedagogical needs. Because of the wide variety of APT attack techniques and scenarios, a comprehensive APTs prevention education game can be quite complex, as it must combine attack techniques with deceiving scenarios. In this paper, we use the frame-based approach to build up the APTs prevention education game, where the frame is used to model the APT attacks scenario and knowledge. Therefore, the properties of the frame model allow us to easily extend or update the APTs prevention scenarios and enrich game contents. To evaluate the effectiveness of proposed approach, we have developed a prototype system to support APT prevention education game. The questionnaire is designed by adopting the definitions in ISO/IEC 9126 to evaluate the performance of our proposed system. The evaluation results show that most of participants are quite satisfied with our proposed system.*

***Keywords*—** advanced persistent threat, frame-based model, APT prevention education game, automatic course generation

I. INTRODUCTION

Among the recent hacking activities, there has been a continuous growth of Advanced Persistent Threats (APTs). APTs are a type of cybercrimes that collect information and data on specific target of enterprise and then perform hacking by using the data. The APTs are more sophisticated than traditional attacks aiming at unspecified targets. It is because once a specific target is selected, information and data can be continuously collected in the form of secret information for a long period of time, and attacks can then be carried out based on the collected information and data. According to the analysis report of Symantec[15], targeted attacks on small companies sharply increased from 18% to 36% or higher. Thus, the systematic and effective handling of an APT by the members of an organization is vital. The APT attack is becoming a serious problem for enterprise and cause heavy financial and political losses[5][6].

APT attacks[1][8], which are more organized and have longer-term political or financial purposes than the traditional attacks made for personal purposes, usually contain a series of scenarios and multiple media, such as web, phone, e-mail, IM, and so on. Attacker applies various scenarios, such as social network web site, e-bank, e-commerce, to defraud victims of critical data and information. This kind of security threats is using social engineering and advanced computer techniques to take advantage of the way humans interact with computers or interpret messages [4]. To educate the active participants of the organization to identify suspicious emails and websites is quite important and is the most efficient way to fight APTs [13]. Since the number of APT cases is continuously increasing, how to manage the APT prevention knowledge structure and transform knowledge into learning materials becomes an important issue.

There are many research results[9][10] showed that game-based learning can increase course satisfaction and learning motivation of learners. And security education games[12][13] can help users to distinguish malicious pages in real world, especially when APTs features nowadays are usually overlooked under the disguise of attractive web contents. However, these education games mainly help users to build the security attack knowledge with limited cases. As to content generation for APT attack, it is difficult to generate various content cases due to the complex and various natures of APT attack scenarios and techniques. And the traditional way is costly and time consuming to create APT cases with a variety of scenarios. Therefore, in this paper, we focus on how to automatically construct more related APTs prevention learning contents according to the new cases.

An APT case[2] is composed of attack knowledge and attack scenarios. The scenario is a series of pages and each page is composed of the APT attack knowledge and APT attack scenario. For example, in a traditional APT attack case as shown in Fig. 1, the APT attack case contains three pages, P01, P02, and P03. The victim accepted a malicious mail in

page P01. Then he/she logined a faking web site with his/her username/password in P02 and P03. Our observation is that attack knowledge and attack scenarios can be simply described by stereotyped attributes. Therefore, in this paper the scenario frame (SF) and the knowledge frame (KF) are used to model the APT attack knowledge and the APT scenario by analyzing the existing APT attack cases. While new APT attack knowledge is obtained, Instantiation property can be used to apply attack knowledge to different scenario; thus a large amount of new scenario pages can be generated. In the experiment, the expert-designed satisfaction questionnaire is used to evaluate the degree of usage satisfaction about the proposed mission management system and the expression capability of the proposed frame hierarchy. Based on the definitions in ISO/IEC 9126[19], we design our questionnaire by considering functionality, reliability, usability, and efficiency. 36 experienced participants were involved in the experiment. The evaluation results showed that most of participants are quite satisfied with our proposed system.

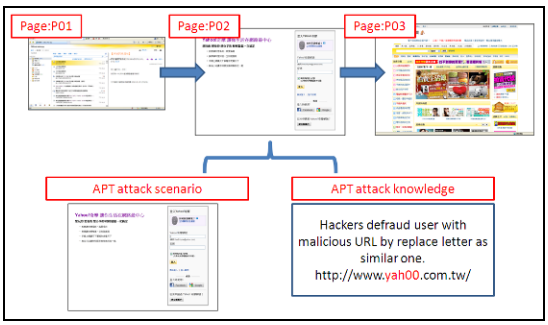


Fig. 1. The Decomposition of a APT attack case

The remainder of this paper is described as follows: Section 2 discusses related works. In Section 3, we propose the frame-based content generation approach and several APT attack techniques. Section 4 presents our APT prevention education game. Section 5 is experimental result and section 6 gives the conclusion and proposes directions for future work.

II. Related Work

A. Advanced Persistence Threat

There are several definitions attempting describe Advanced Persistent Threat (APT). Advanced persistent threat (APT)[1] usually refers to a group or an individual hacker with both the capability and the intent to persistently and effectively target a specific and valued entity. APT uses a series of hacking techniques to trick users into giving away sensitive information, such as phishing, infected media, supply chain compromise, and social engineering. Generally, persistence and targeted attack are characteristics of APT. In the last few years, several incidents illustrated the criticality of targeted attacks. In which, emails or other forms of electronic communication are sent to users that purpose to come from trusted sources, such as their bank or company IT administrator. According to Anti-Phishing Working Group (APWG) in a report issued in 2011, 37% of respondents to its survey reported that they had phishing or spoofed sites planted on their web server two or more times in the previous year, which the APWG states points to the increasing persistence of hackers. A particular trend being seen by the APWG is the use of spear-phishing and this is something that has been widely used in the APT attacks. [17][21][22][23][24]

B. Security Education for APTs

Education is one effective way to protect users from attacks on Internet. There are many researches focused on APT prevention educations and anti-phishing that can be classified into educational documents and education games. Robila et. al. [20] used Phishing IQ Test and class discussions to educate students. Several APT education games have been proposed for improving the learning motivation of users [3][7][11][14][18]. However, some studies indicated that security educational materials are not sufficient for anti-phishing and APT prevention. The primary reason is that users usually have low motivation of studying documental anti-phishing materials. Furthermore, without realistic simulation, the non-simulation games contain only general principles or limited security attacks knowledge (such as basic URL obfuscations). Compared to various advanced attack techniques, users may not easy to learn these attacking features, resulting in the difficulty of applying APT knowledge in real environment. For example, there is no visual difference between legitimate and fake URLs when applying graph substitution attack utilizing a graph of legitimate URL to redirect users into malicious site.

In transditional simulating games, it is usually costly and time consuming to update new phishing case in education material due to the internal fixed content templates are not flexible enough to generate learning pages for different page scenarios and attack knowledge. Moreover, the variants of the given scenarios in web page are required to be provided to test whether users can apply APT attack knowledge to real environment. Thus, our idea in this paper is to find a way to automatically generate game content; e.g., extend items to generate new phishing pages. The issue of game-based APT prevention education is how to generate material cases into game missions as learning materials.

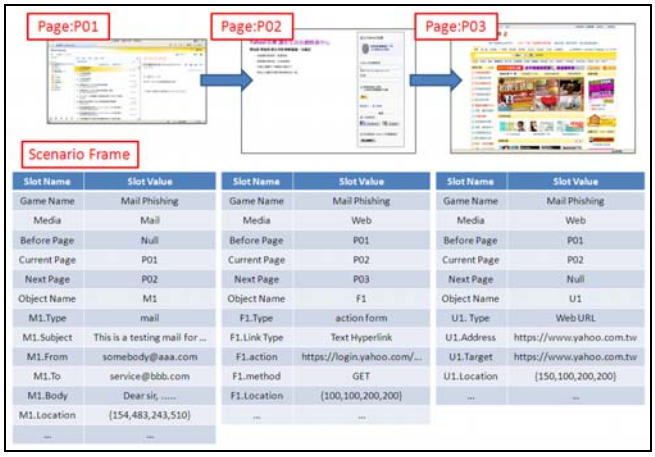
III. Frame-based Content Generation Approach for APT Scenario Modeling

The goal of APT prevention education game is to develop the users' skill of discerning malicious attacks in the real world. APT attacks consisting of a series of malicious scenarios can be divided into two parts, APT attack scenario and APT attack knowledge. In the paper, we propose the frame-based content generation approach to describe the scenario. Then, we apply the frame-based approach to support the APT prevention education game construction. Our frame-based approach contains two main parts, APT attack scenario frame (SF) is used to represent APT attack scenarios and APT

attack knowledge frame (KF) is used to manage APT attack techniques.

A. APT attack scenario frame (SF)

In this sub-section, an APT attack scenario frame (SF) is created by the common APT case with multiple scenes. In Fig.3, the scenario consists of three scenario pages, P01, P02, and P03. Attacker defrauds victim with malicious e-mail and interesting topic in P01. Then, users log on the web site with obfuscated URL by keying in username and password in P02. Finally, hacker acquires information of username and password and then redirect user to real web site in P03. Once users overlook the legitimacy of visiting pages or hyperlinks, their sensitive private information is exposed to the risk of being leaked. Such kind of attack is called "URL obfuscation". The objects representing the items in scenario page are also defined in a scenario frame. In this case, the objects M1, F1, and U1 are represented as frames. We transform the mail, web action form, and URL address as objects, M1, F1, and U1, respectively.



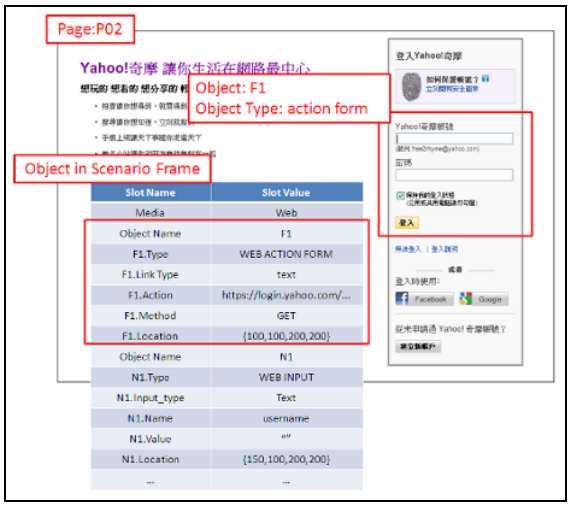
Slot Name	Slot Value
Game Name	Mail Phishing
Media	Mail
Before Page	Null
Current Page	P01
Next Page	P02
Object Name	M1
M1.Type	mail
M1.Subject	This is a testing mail for ...
M1.From	somebody@aaa.com
M1.To	service@bbb.com
M1.Body	Dear sir, ...
M1.Location	(154,483,243,510)
...	...

Slot Name	Slot Value
Game Name	Mail Phishing
Media	Web
Before Page	P01
Current Page	P02
Next Page	P03
Object Name	F1
F1.Type	action form
F1.Link Type	Text Hyperlink
F1.action	https://login.yahoo.com/...
F1.method	GET
F1.Location	(100,100,200,200)
...	...

Slot Name	Slot Value
Game Name	Mail Phishing
Media	Web
Before Page	P02
Current Page	P03
Next Page	Null
Object Name	U1
U1. Type	Web URL
U1.Address	https://www.yahoo.com.tw
U1.Target	https://www.yahoo.com.tw
U1.Location	(150,100,200,200)
...	...

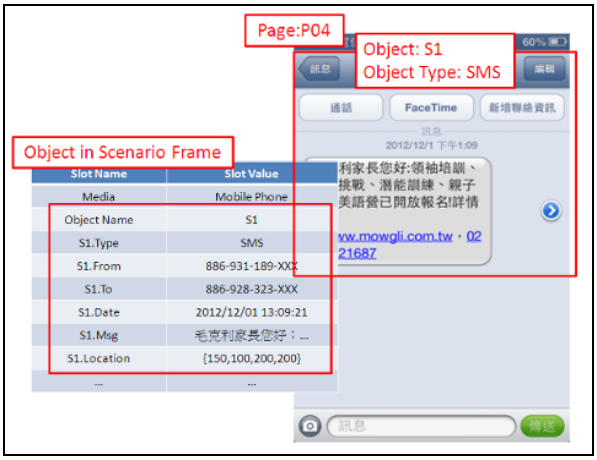
Fig. 2. Frame-based of Education Scenario for APT Prevention Game

To represent the characteristics of the objects embedded in scenario page, each type of object has its own specific attributes. As shown in Fig.3, the object F1 represents the web action form with the frame attributes including type, link type, action, method, and location. In Fig.4, the example object S1 represents the SMS of mobile device with attributes of object S1 including type, from number, to number, date, message, and location.



Slot Name	Slot Value
Media	Web
Object Name	F1
F1.Type	WEB ACTION FORM
F1.Link Type	text
F1.Action	https://login.yahoo.com/...
F1.Method	GET
F1.Location	(100,100,200,200)
Object Name	N1
N1.Type	WEB INPUT
N1.Input_type	Text
N1.Name	username
N1.Value	""
N1.Location	(150,100,200,200)
...	...

Fig. 3. An example of object F1: action form in Web application



Slot Name	Slot Value
Media	Mobile Phone
Object Name	S1
S1.Type	SMS
S1.From	886-931-189-XXX
S1.To	886-928-323-XXX
S1.Date	2012/12/01 13:09:21
S1.Msg	毛克利家長您好：...
S1.Location	(150,100,200,200)
...	...

Fig. 4. An example of object S1: SMS message in mobile phone

B. APT attack knowledge frame (KF)

In our frame-based approach, we acquire APT attack knowledge from domain experts to construct the knowledge frame (KF). Below is an example of the APT attack knowledge frame (KF). The attack knowledge can be represented as frame with six stereotype attributes of the APT attack knowledge frame. "Security Issue" is the title of attack. "Security Issue No" gives the knowledge number. "Media" stands for the container of APT attack, such as email, web, and mobile. "Suit To" represents the scenario pages that are associated to attack knowledge. "Issue Description" describes the attack with a short description. "Action Script" is the script language used for object operation. We can use action script to manipulate the object value in scenario pages. Fig.6 shows the slot attributes of the attack knowledge frame describing the case of "URL obfuscation". This APT attack replaces the letter by similar one in page content appearance as well as replacing URL appearance to trick the users. For example, in knowledge K026 and K027, the letter "O" can be replaced by zero and the letter "l" can be replaced by one.

Slot Name	Slot Value	Slot Name	Slot Value
Security Issue	Homography with similar word (o, O, 0)	Security Issue	Homography with similar word (l, I, 1)
Security Issue No	K026	Security Issue No	K027
Media	Web, Mobile, Mail	Media	Web, Mobile, Mail
Suit To	S01, S02, S03, ...	Suit To	S01, S02, S03, ...
Issue Description	Homography by replace with similar word.	Issue Description	Homography by replace with similar word.
Action Script	If Object.String contain "o" then replace with "0". If Object.String contain "O" then replace with "0". If Object.String contain "0" then replace with "o". If Object.String contain "0" then replace with "O".	Action Script	If Object.String contain "l" then replace with "1". If Object.String contain "I" then replace with "1". If Object.String contain "1" then replace with "l". If Object.String contain "1" then replace with "I".
...	APT attack knowledge frame (KF)		...

Fig. 5. Examples of APT attack knowledge frame

C. APT scenario extensions with knowledge frame

Based upon the features of frame-based approach, the new APT attack knowledge can be produced easily with frame format, and such a frame-based format easily makes the knowledge extension. There are two extension methods. First, we can extend the malicious APT attack scenario content by combining the scenario frame (SF) and knowledge frame (KF). In Fig.6, we find out that the F1.action in scenario frame contains URL "http://www.yahoo.com.tw". According to action script in KF, we can modify the URL by replacing similar letter(s) in URL address. A new scenario content can then be generated with a malicious F1.action URL "http://www.yah00.com.tw". This is a "URL obfuscation" attack. Users have different degrees of alertness on different APT scenarios even when they are implemented by the same APT attacking techniques. For a given APT attack knowledge, we can apply it on all the scenario pages satisfying the precondition of APT attack knowledge. For example, the APT attack knowledge in Fig.6 can also be applied in other scenario page, such as mobile and email, and hence can generate other scenario content. This instantiation property provides us another way to enhance the extensibility of game contents by the inheritance property.

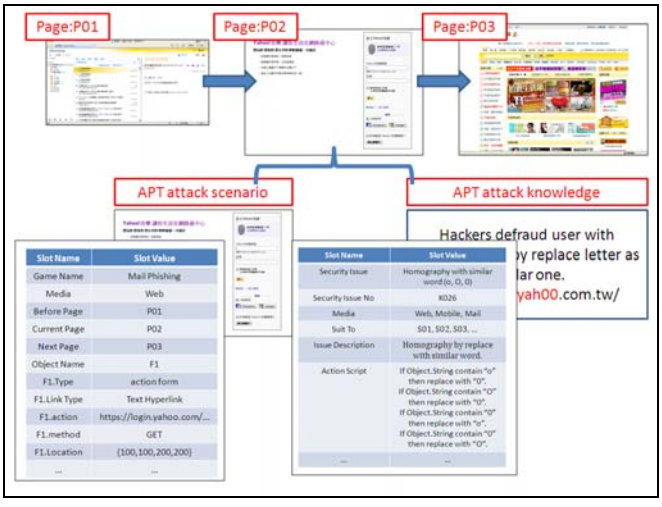


Fig. 6. Scenario content generation by combining SF and KF

Secondly, we combine different kinds of attack knowledge (KF) to generate new attack knowledge (KF). With the advantage of frame-based architecture, experts can easily combine APT attack knowledge, this process is called knowledge fusion. Once the two APT attack knowledge frames(SF) intersect, they can be fused into a new APT attack knowledge frame. It can be considered as the new APT attack knowledge implemented by these two APT attack techniques. For example, "Homography with similar word (o, O, 0)" can be used to create a spoofed URL by replace O in URL by 0; and "Graph substitution attack" is to maintain the same link appearance by embedding graph file and the corresponding underlying link target that leads to different URL of APT scenario. These two APT attack techniques can be fused into "Graph substitution attack with similar character attack on link target: O and 0".

IV. APT PREVENTION EDUCATION GAME

Based on the frame-based model, we proposed our prototype system of APT prevention education game. The architecture of prototype system is shown in Fig.7. There are three main parts in our system. 1) In editor interface module, we propose authoring tools for the domain experts to edit the structure of scenario pages. The system assists experts to generate testing content automatically by providing APT attack scenario and APT attack knowledge. 2) In content generation module, the APT attack knowledge and APT attack scenarios are maintained in knowledge frame repository and scenario frame repository, respectively. 3) In user interface module, our prototype system is web-based and users can access APT prevention education game via standard browser. The inference engine applies the APT scenario contents and strategies based on user profiles.

There are two roles in our prototype system. APT prevention education game can describe nowadays APT attacks via frame-based content. We proposed authoring tools to satisfy domain experts' requirement in content generation. Furthermore, the APT prevention education game can attract users' attentions. In Fig.8, domain experts can edit new scenario page for specific APT attack technique by combining attack scenario and attack knowledge. Therefore, APT prevention education game can be flexibly for adapted for generating constantly evolving APT attack knowledge. In Fig.9, users can play the APT prevention game by accessing the web-site. To simulate the general experiences of Internet usage, users need to complete some missions. In mission, users can determine which part in screen is malicious attack.

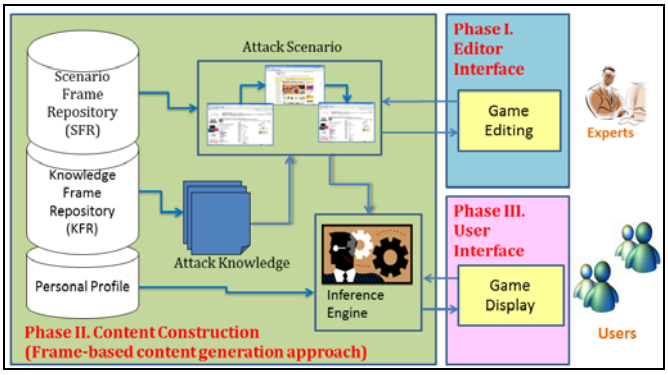


Fig. 7. The system architecture of APT prevention education game

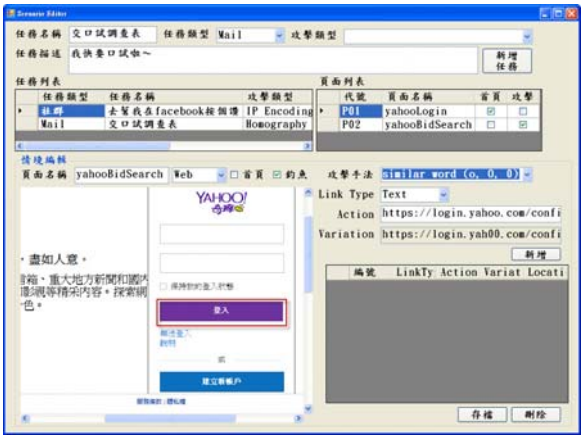


Fig. 8. Screen shot of editor interface of APT prevention education game

Fig. 9. Screen shot of user interface of APT prevention education game

V. Experiment Results

To evaluate our education game, we request 36 students majoring computer science in the university in Taiwan to participate in our experiment. In the experiment, a test composed of 20 APT cases is given and participants were asked to determine whether each page is APT attack or not and the confidence of their judgments. After a training session of the APT prevention education game, another 20 APT cases are given. Finally, a questionnaire based on definitions of standard ISO/IEC 9126 was applied for evaluating the performance of our education game. ISO/IEC 9126 is an international standard for evaluating the software quality. The fundamental objective of this standard is to address some of the well-known human biases that can adversely affect the delivery and perception of a software development project. In Table 1, we evaluate our system with 12 questions based on the definitions in ISO/IEC 9126 with four factors, functionality, reliability, usability, and efficiency. We evaluate the students' learning satisfaction by a questionnaire. After the participants finished APT prevention education game, they were asked to fill a satisfaction questionnaire as well as providing any comments. Five-point Likert scale is used to evaluate the degree of satisfaction, ranging from strongly disagree (1) to strongly agree (5). The results from the questionnaire are shown in Table 1.

In Table 1, the results of questionnaire are used to measure the performance of the proposed education game in four factors. We found out that the average satisfaction score (column Avg. Sat.) of usability is higher than other factors. A reasonable speculation is that the proposed frame-based mechanism allows us to control the flow of scenario easily, which greatly enhances the game user-friendliness. Q7 and Q8 measure perceived ease-of-use in the game. The results show that the participants believe the interface of education game is friendly. In the meantime, the reliability and efficiency get lower satisfaction scores. The average overall satisfaction is 3.86, which indicates that our APT prevention game is helpful and easy to use.

TABLE I
System Satisfaction for Participants

Category	Question	Sat.	Avg. Sat.
Functionality	Q1.I can aware APT attack patterns by playing the game	3.85	3.783
	Q2.The hint of game can help me learn how to detect the phishing websites	3.83	
	Q3.The game can help me learn how to detect the phishing websites	3.71	
	Q4.I can learn how to use APT prevention tools in the game	3.74	
Reliability	Q5.The fault-tolerance mechanism of education game is well.	3.69	3.605
	Q6.I can rollback or redo the process easily in game operating.	3.52	
Usability	Q7. The game control is easy to play	4.17	4.193
	Q8. The proposed scenarios of game are similar to real world application.	4.03	
	Q9.I will play the game if I need to know APT prevention knowledge	4.26	
	Q10. The guidance can help me understand how to play the game	4.31	
Efficiency	Q11.The game learning is efficient.	3.21	3.425
	Q12.The education game can reduce learning time and enhance learning efficiency.	3.64	
Summary		3.86	

VI. Conclusion

The main contribution of this paper is to model the APT scenario with knowledge frames and scenario frames, because APT scenario page can be decomposed into APT attack

knowledge and APT attack scenario. Then, the relations between APT attack knowledge and scenario pages can be modeled by the Instantiation property, and the revisions of different APT attack knowledge can be easily done based upon the properties of frame model. From our evaluation, the results showed that the proposed APT prevention education game is effective for describing nowadays APT attack knowledge and educating users how to distinguish APT attacks in real life.

In the near future, we will continue to increase game elements of APT prevention education game for the purpose of attracting users' attention. In the meanwhile, we will also develop new adaptive learning techniques to shorten the time of users studying APT attack knowledge.

ACKNOWLEDGMENT

This work was partially supported by the National Science Council of the Republic of China under grants NSC101-2511-S-468-007-MY3 and NSC102-2511-S-468-003-MY2.

REFERENCES

[1] "Advanced persistent threat Wikipedia", http:// http://en.wikipedia.org/wiki/Advanced_persistent_threat/.

[2] A. J. Ferguson, "Fostering E-Mail Security Awareness: The West Point Carronade,". EDUCAUSE Quarterly, 28(1), 54-57.

[3] A. K. Krishna, A. N. Kumar, "A problem generator to learn expression evaluation in CS I, and its effectiveness," Journal of Computing in Small Colleges. 16,4

[4] B. Schneier, "Semantic attacks: The third wave of network attacks, "Crypto-Gram Newsletter. http://www.schneier.com/crypto-gram- 0010.html#1, 2000

[5] Chen, T.M.: Stuxnet, the real start of cyber warfare? IEEE Network (2010)

[6] Farwell, J.P., Rohozinski, R.: Stuxnet and the Future of Cyber War. Survival: Global Politics and Strategy (2011)

[7] I. A. Zualkernan, S. A. EI-Naai, M. Papadopoulos, B. AI-Amoud and C. Mattews, "Automatic Generation of Just-in-time Online Assessments from Software Design Models," Educational Technology & Society, 12 (1), p. 173–192.

[8] Jeun, Inkyung, Youngsook Lee, and Dongho Won. "A Practical Study on Advanced Persistent Threats." Computer Applications for Security, Control and System Engineering (2012): 144-152.

[9] J. W. Chen, "Designing a Web-based van Hiele Model for Teaching and Learning," Computer Programming to Promote Collaborative Learning, the Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05, 2005.

[10] L. Pannese, and M. Carlesi, "Games and learning come together to maximise effectiveness: The challenge of bridging the gap," British Journal of Educational Technology, 38(3).

[11] P. Deane, and S. Kathleen, "Automatic Item Generation via Frame Semantics : Natural Language Generation of Math Word Problems," in Annual Meetiing of the National Council on Measurement in Education.

[12] P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Protecting people from phishing: the design and evaluation of an embedded training email system," In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07).

[13] P. Kumaraguru, S. Sheng, A. Acquisti, L. F. Cranor and J. Hong, "Teaching Johnny not to fall for phish," ACM Transaction on Internet Technology, 10(2).

[14] S. Kim, J. A. Clark and J. A. McDermid, "The regorous generation of Java mutation using HAZOP," in In Proceedings of the 12 the International Conference on Software and Systems Engineering and Their Applications (ICSSEA '99).

[15] Symantec, Symantec report finds that more than a third of global targeted attacks are aimed against small businesses. Symantec Press Release (July 20, 2012) http://www.symantec.com/about/news/release/article.jsp?prid=20120710_01

[16] S. S. Tseng, K. Y. Chen, T. J. Lee, J. F. Weng, "Automatic content generation for anti-phishing education game," IEEE International Conference on Electrical and Control Engineering (ICECE), pp. 6390-6394, 2011.

[17] Aloul, Fadi, Syed Zahidi, and Wassim El-Hajj. "Two factor authentication using mobile phones." Computer Systems and Applications, 2009. AICCSA 2009. IEEE/ACS International Conference on. IEEE, 2009.

[18] Z. Branko, S. Stankov, M. Rosic and A. Grubisic, "Dynamic test generation over ontology-based knowledge representation in authoring shell," Expert Systems with Application, 36, p. 8185-8196.

[19] "ISO/IEC 9126" http://en.wikipedia.org/wiki/ISO/IEC_9126.

[20] Robila, S. A., and Ragucci, J. W, "Don't be a phish: steps in user education," In ACM SIGCSE Bulletin, Vol. 38, No. 3, pp. 237-241, June, 2006.

[21] Jeun, Inkyung, Youngsook Lee, and Dongho Won. "A Practical Study on Advanced Persistent Threats." Computer Applications for Security, Control and System Engineering. Springer Berlin Heidelberg, 2012. 144-152.

[22] Baize, Eric. "Developing secure products in the age of advanced persistent threats." IEEE Security & Privacy 10.3 (2012): 0088-92.

[23] Li, Frankie, Anthony Lai, and Ddl Ddl. "Evidence of Advanced Persistent Threat: A case study of malware for political espionage." Malicious and Unwanted Software (MALWARE), 2011 6th International Conference on. IEEE, 2011.

[24] Tankard, Colin. "Advanced Persistent threats and how to monitor and deter them." Network security 2011.8 (2011): 16-19.

Using Forensic Images to Teach About White Collar Crime Online

Professor and Cybercrime Training Lab Director Eamon P. Doherty CCE, CPP, SSCP
[1]School of Administrative Science, Fairleigh Dickinson University, Teaneck, N.J., USA
Doherty@fdu.edu, FDU, 1000 River Road, H-DH2-16, Teaneck, NJ, 07666
Ester Doherty and Elly Goei, FDU Alumni

Conference : EEE2014

Abstract - This paper discusses the concept of using forensic images and low cost e-discovery tools for helping students learn the practical component of a white collar crime class. Since white collar crime is often abstract and difficult to define, having a practical component of the class in tandem with traditional printed materials should help students grasp the concepts. Forensic images include all the allocated and unallocated space on a drive. It is possible for some private communications or data to be released on a forensic image unless every byte is reviewed by the educator. This paper suggests using the CFReDS forensic images available from NIST in white collar crime classes. Lastly, the paper discusses the necessity of e-Learning as the most viable option for many students who are law enforcement and security professionals that have rotating shifts and unpredictable schedules.

Keywords: White Collar Crime, e-Forensics, Online Learning

1 Introduction

Students often take theoretical classes about white collar crime and learn many concepts but do not get an opportunity to apply that knowledge. This paper discusses a way of teaching students about forensically imaging a machine with simulated white collar crime evidence on it, creating an iso image of the entire hard drive, and then examining that image in Oracle's VM VirtualBox 4.1. The result is that students can open up a window on their desktop and have what appears to be the simulated white collar criminal's computer running on their desktop. The experience is beneficial because students learn about creating a forensic image, namely an iso image of a hard drive, and then looking through the image. Creating a forensic image of a suspect's drive is considered the most important task in computer forensics [1]. Some people who are certified computer examiners (CCE) make a career of forensically imaging laptops, desktops, notebooks, and hard drives. Loading the iso image of the drive into a program called VirtualBox is also important because it teaches the lesson of creating a virtual machine. Virtual machines are very important because one can run another instance of a computer in a window without causing any permanent damage to the host machine or to the preserved evidence. When the virtual machine is extinguished, the iso image is not changed. The totality of creating a forensic image of a suspect's computer, creating a virtual machine, and viewing it for evidence are important skills for anyone wishing to enter the field of eDiscovery or eforensics. The skills are also good for anyone who is seeking to work in the white collar crime unit of a county prosecutor's office.
Some students and researches may want to see how an application may behave on an older operating system that they do not have installed on their computer. VMWare Player can be used for creating another operating system on the examiner's computer. An examiner may have Windows 7 but might need to run an application in Windows XP or Windows 98 to see how it would behave. VMWare Player and an old iso of Windows 98 could be installed on the Windows 7 machine so that the student or examiner does not need to purchase another computer for the experiment.

2. PC Hardware & Operating Systems

Many places complain about costs and wanting to do everything at little cost. It would be good to use machines that were going to be discarded. One would remove all the personal data and documents and then use a tool such as CCLeaner to remove those documents and wipe the free space and temp files where other data might be. Then a computer with a 1 GB or 2GB hard drive with Windows 98 would be a good choice. A newer system would be too large and require the student to look through too many files. The system just described could have a forensic image of the 1 GB hard drive burned in an ISO format to a DVD. A 1 GB ISO image on a DVD could be mailed to students or might possibly be downloaded from an online learning environment. An image of a newer system would be too large and the downloading time would be too long and the data transfer would cost the school too much money. Older computers are often available at companies and schools for little or no cost.

3. Creating Simulated Criminal Data

You could have two computers with Microsoft Outlook on them. Then both people could email each other about a simulated car repair franchise. Person one could pretend to be a manager ordering person two to break things on incoming cars in order to make money on repairs. Person one could also tell person two to pressure each customer into getting an alignment and replace some slightly worn parts. Person two could have email conversations with coworkers about feeling guilty for doing unnecessary work and taking money from senior citizens on a fixed income. Person one could also have two sets of books in MS Excel consisting of one for himself and the other to show the IRS and his accountant. The second set of books contains work that was paid by check or credit card while book one has the same information as well as all the orders paid in cash. The books would show that he did not report all his income and evaded taxes. The computer could also have email between the company and local law enforcement showing some "campaign contributions" in exchange for losing some complaints about the car business. The computers would have emails in the .pst files and .xls MS Excel files containing the fraudulent business transactions.

4. Creating the ISO Image

The next step would be to connect a USB external drive and use a program such as IMGBurn or MagicISOmaker http://www.magiciso.com/ to create an image of the simulated suspects' hard drives. Once the ISO image is on the external hard drive, it can later be burned to a DVD or loaded in an online learning environment for students to download. There are many tools online that allow one to create an ISO image. The concern about creating a forensic image is that there may be an existing file or remnants of a deleted file that contain some private communication or data should not be released to the public. One could not create a forensic image of a machine that was connected to the Internet and used by students in the lab because the forensic image might contain usernames, passwords, internet browsing histories, and private communications that would be embarrassing for the student. Suppose a student reported something confidentially and this was revealed because of some text found in an image. Such a release could damage the reputation of the school and perhaps lead to a lawsuit. Even if the lab with student computers uses a standard image with scrub program that deletes all the activity from the previous session, who is going to examine every byte to see that there is nothing private in hundreds of gigabytes of a forensic image ?

It would therefore be more prudent to use a forensic image made by someone else. The National Institute of Standards and Technology has forensic images in dd files or Encase format that can be used for examination in class. This is not only timesaving but removes the possibility of accidentally releasing private information that exists on a computer hard drive.

5. Using Computer Forensics Reference Data Sets CFReDS

The National Institute of Standards and Technology is in the process of creating a set of forensic images for use with computer forensic tools [5]. This process is being funded by the National Institute of Justice and is called the "Computer Forensics Reference Data Sets [5]." The NIST CFReDS website states, "Investigators could use CFReDS in several ways including validating the software tools used in their investigations, equipment check out, training investigators, and proficiency testing of investigators as part of laboratory accreditation. The CFReDS site is a repository of images [5]." As educators and students, we find this exciting because there is a downloadable forensic image, a set of questions, and a controlled release of answers that we can use to test our proficiency as examiners as well as the effectiveness of our tools. There are a variety of cases, questions, and forensic images that can be downloaded and investigated. These data sets could also be used in an online class to help students validate their investigation tools. These forensic images could also be utilized by students to provide a metric to measure their skills as investigators.

6. Learning About White Collar Crime

There are many good books on this subject but two books that stand out. The first book is "Trusted Criminals: White Collar Crime In Contemporary Society" [2]. The second book is "The Criminal Elite: Understanding White Collar Crime. [3]" It is important to have easy to read books that give specific legislation and examples of white collar crime. These books are also good because they also discuss the problem of using too broad a definition of white collar crime as well as using a definition of too narrowly defining a white collar crime. It is also important to remember that classes contain groups of international students from non western developing nations and what might be assumed to be white collar crime by some may be business as usual for others. Therefore it is important to define terms, explain concepts, and not assume that there exists a common culture or set of experiences among the students.

7. E-Discovery Tools

It is also important for students to get used to using eDiscovery tools which may be utilized by small law firms and computer forensic companies who may hire them. Many of the eDiscovery tools presently used by large law firms are very expensive and not cost effective for schools to purchase for teaching purposes. They are also too expensive for small law firms who must handle small cases of people with 800 GB hard drives [4]. Tom O'Connor, director of the Gulf Coast Legal Technology Center in New Orleans, Louisiana, has been addressing the problem of locating lower cost eDiscovery tools for small law offices.

Tom O'Connor has compiled a list of software packages that are cost effective for small law firms who handle small cases [4]. The authors of this Worldcomp paper also hold the opinion that these lower cost tools could be purchased for schools and used by students who are studying white collar crime. Quickview Plus was one of the tools that he recommended and was one we had experience with since it was made available with the Certified Computer Examiner (CCE) Toolkit [4]. It is an excellent tool for looking at documents and email that are in common or obscure data formats. This would be a good teaching tool because it is relatively inexpensive, has a point and click interface, and might also be a tool that students may see again in a small law firm.

The tool dtSearch Desktop is another tool that Tom O'Connor recommends for small firms working on small cases. This is another tool that many CCEs use on the job and is useful for finding various strings of words or individual words that may be located in numerous documents on a drive [4]. Tools such was dtSearch and Quickview Plus are also good because they are more intuitive and do not require expensive training classes and certification. Tom O'Connor also suggests Acrobat Legal Edition, Safe Copy, and Harvester [4]. He also gives the link for obtaining more information. We three authors of this paper have no previous experience with these tools. However; from what the websites show, they appear to be good candidates for purchase and use for teaching forensics. They would also appear to be useful in examining simulated evidence in a white collar crime class.

Winhex by X-ways was one of the tools that was not on the list but should be considered. There is also a limited function free version that can be downloaded and used in a class. The paid version of WinHex lets students and examiners do a forensic image of an entire drive. That means all the allocated and unallocated space. That means that a copy of every bit of unused and used space on a drive could be copied. Winhex also lets the examiner search used and unused space on the drive. This is important if someone deleted something and it was still there while the space was waiting to be written over. X-ways also has many forums and blogs online so it is possible for students to post questions and get answers from experts in the field. The experience from such networking is also invaluable. Such connections could be useful for future technical questions or to learn about entry level jobs.

8. Storage Area Network (SAN)

It is also good to have a storage area network in a school so that students could learn to cooperate on cases at their workstation in a classroom. Storage Area Networks are often called SANs and are used by many law enforcement agencies and private computer forensic agencies. If the class was done online, a website could be used and students could share a file in a directory. Students who are online could also keep in contact with each other with videoconference tools such as Uvu or Skype. We used Uvu with a Radio Amateur Civil Emergency (RACES) group and six people could easily videoconference on Smartphones at no cost. GoogleDocs or DropBox are other tools that could be used in an online class for collaboration. There are many tools for video conferencing and sharing documents. Teachers familiar with technology have a range of tools to choose from.

9. Webinars and MOOCs Help People with Unpredictable Schedules

Massive Online Open Courses are known as a MOOC. They are becoming quite popular worldwide for teaching technology. The three authors of this paper often go to ASIS International meetings for luncheons and to learn something from a guest speaker. Some of the attendees are forensic experts and investigative specialists who often get called at a moment's notice to investigate security conditions at a property where a severe injury or death occurred. These individuals appear to have sufficient ability and funds to get training but the nature of their job does not allow a predictable schedule to take classes. These individuals have told us that free MOOCs and webinars are convenient for them. If they can make it, they participate online. If they cannot go, nothing is lost and the presenters of the webinars are said to be generally understanding of such circumstances. Many security professionals told us that any type of classes that can be selected online and taken for free are also great. FEMA uses this model of learning and has a large selection of classes on a variety of topics for the emergency management community. Each course ends with a test. If a certain grade is achieved, it shows some proficiency and a certificate can be printed. It seems that MOOCs, webinars, and various modes of asynchronous learning work well for people with rotating shifts or unpredictable schedules.

10. Books, Journals, e-magazines, and e-Journals

There are a variety of online and paper based materials about the subject of white collar crime and investigating computers. It is good to encourage students to read these to see what areas of investigation that they are interested in. It is also good for students to look at the credentials of the authors and see what certifications, experiences, and education that they have. If a certain certification such as the CCE or Encase certification often appears, then it would behoove the student to seek obtaining that certification. E-magazines are often free or low cost and allow the student to see recent articles that may suggest a new lucrative trend in eForensics or investigating white collar crime. Many of these books, magazines, and journals also have contact information for the author and it may be good for students to write a note stating that they are new to the field and are seeking advice. Encouraging students to use reliable online libraries are important because they need to get in the habit of obtaining reliable information. If they become computer forensic investigators of white collar crime, they could be discredited if an investigation ever went to court and they were found to have used unreliable sources.

Many universities such as FDU also have holdings of over twenty thousand e-books, e-journals, and e-magazines. An FDU university student can do tremendous amounts of research from home during non-traditional hours. Many other universities and community libraries have similar low cost or no cost options that students can take advantage of.

11. Conclusion

The Webinars, MOOCs, and various forms of online learning are important for those people who wish to learn about white collar crime but have a profession that does not have a predictable schedule and allow them to schedule traditional in person classes. An online class with links to forensic tools, a forensic image, and educational material holds the most promise for teaching busy professionals about investigating white collar crime on a computer. Using a tool such as GoogleDocs might be a good way to simulate a storage area network (SAN) and have students work collaboratively on a case. It might also be a good idea to have the students use low cost eDiscovery tools such as dtSearch Desktop, Quickview Plus, and WinHex because they might be used in a forensic investigation business or small law office. It would also be a good idea to have students create a virtual machine with tools such as VirtualBox and learn to load iso images so that they will have some experiences that are valued by industry. Lastly, it would be good to encourage students to become certified in investigating computers and white collar crime.

12. References

[1]. Doherty, E. (2013), "New Challenges in Teaching e-Forensics Online", WorldComp 13, EEE 2013 Proceedings, ISBN 1-60132-240-2, P67

[2]. Friedrichs, D., (2009), "Trusted Criminals: White Collar Crime In Contemporary Society" Cengage Learning, Printed in the USA, ISBN-10: 0495600822, 496 pp

[3]. Coleman, J., (2006), "The Criminal Elite: Understanding White-Collar Crime" Publisher: Worth Publishers, ISB: 0716787342

[4]. O'Connor, T, (2013)" Cost Effective E-Discovery For Small Cases" Volume 30 Number 1, American Bar Association http://www.americanbar.org/publications/gp_solo/2013/january_february/cost_effective_ediscovery_small_cases.html

[5]. URL Accessed 1/22/2014, The Computer Forensic Reference Data Sets (CFReDS) Project, http://www.cfreds.nist.gov

The Cyber Education - Security for Citizenship

A. Joel Mana Gonçalves[1], and B. Rodrigo Stábile[2]
[1]Directorate of Services to the Citizen, Prodesp - Cia. de Proc. De Dados do Estado de São Paulo, Taboão da Serra, SP, Brazil
[2]Legal Advice Management, Prodesp - Cia. de Proc. De Dados do Estado de São Paulo, Taboão da Serra, SP, Brazil

Abstract *- This This paper considers the new world scenario caused by the benefit of the virtual world, massively distributed via the Internet, which quickly became the information in the most valuable asset in the beginning of this millennium. Information is a vital asset which is present in all social strata of the modern world. This information, that is everywhere, has been gaining special value in the spheres of government and international relations. And before this present scenario as citizenship wins aspects to be exploited and disseminated as basic design information. We also commented that a well-informed and culturally educated citizen will have better opportunities, and improve their critical sense, their formal and emotional intelligence, and other benefits.*

Keywords: *education, security, government, citizenship*

1. Introduction

There is a new world scenario experienced by the benefit of the virtual world. This panorama, diffused by the massive use internet, quickly imposed at all, honoring the “information" as the most valuable asset of the beginning of the current millennium.

The " information”, as an indispensable well, is present in any social stratum of the modern world since the start of the life of individuals, however, it is undeniable that earns more emphasis and value on the spheres of government and international relations between states, it serves to guide the conduct and behavior of all methods of strategic positions in government in various segments.

Contrary to much spreads nowadays, is not from today that access to information is considered paramount, it always has been actually. However, the speed of the information traffic gained in platform called "internet”, made it the darling of today, and the great tool to be exploited by all, even by enthusiasts of technologies.

In fact, information in its many facets : written, spoken (videos) , pictures (even that encrypted), to conform to the platform of the internet and its speed, remained prominent for providing the concept of "online", or "full time" . However, just as an example, long television broadcasts values (or a program) made live, what you might consider embryonic these new times.

But precisely because of modern times, it seems we are always outdated on a given fact not come to our attention - which is horrible! As is also true , another bias, we feel the slightest of the mortals when comment on something that our neighbors still have not learned , feeling it real " dinosaur" that need to quickly return to planet Earth .

In this respect, even the very reputable teaching of Brazilian lawyer Dr. Patricia Peck explaining the new times:

> “A little over forty years ago, the Internet was just a project, the term “globalization” had not been coined and data transmission by optical fiber does not exist ... With the changes since we entered the era of real-time the virtual displacement of the business, breaking paradigms. This new era brings transformation in various segments of the society ... The law is also influenced by this new reality. The dynamics of the information age requires a deeper change in the very way the Right is exercised and thought into their practice"

This point and already talking about a little virtual education and digital security is salutary the comment to the doctrine of the prestigious professor at the University of São Paulo Celso Antonio Pacheco Fiorillo and of the teacher Christiany Pegorari Conte that notes :

> “Infeasible discuss the advent of the Information Society without putting in a prominent position the Internet as well as

> their reflections on the legal reality of the community. As, especially the Internet, information and communication technologies have brought the need for a new look at old rights, such as the information, communication, freedom of expression and to privacy as well as the question about the emergence of new goods require a specific legal protection (as in the case of so-called Computer Security, which covers the completeness of the information posted on the worldwide web, the availability of access and confidentiality of the information) p.12.
>
> The law must adapt to the new reality, under penalty of losing his true role, namely disciplinary social relations and enforce standards of conduct. Thus, the binomial law and does not constitute Internet phenomenon passenger. It is a reality not explored, but that should be studied in all fields of legal sciences, in order to secure new rights and the enforcement of existing ... We live in a society marked by called Digital Revolution. Concepts such as the Internet, global village, virtual space and eliminating boundaries mark the social reality of the twenty-first century. In this context of virtual reality, new relationships are consolidated at each instant, requiring this way, legal guardianship, to ensure effectiveness and safety of such relationships p.15."

Who have not come across a video, a written document, or a review of social network that quickly gained countless levels and frightening viewing and comments? Indeed, just as a comment, in this endeavor the nuisance actions have far more attractive valuable human minds and, often, are the most striking.

Despite it being so when the Internet did not exist , the fact is that something negative always drew more attention than the politically correct attitudes, behold, good examples almost always are seen as a liability .

Against this background it is not necessary to conclude that the " digital education toward citizenship ", the purpose of this text, also won highlighted the difficulty and, at this level, concepts of citizenship may also surface in the virtual environment , as the internet has the capacity to make them, though virtual, human relations gain even greater depth and intensity , eliminating the physical proximity but honoring the essence of pleasure be valued in the other, even if a screen, a click on a " tanned ", a "re-tweeting" etc. .

And it is precisely this scenario that citizenship (transmitted in digital and/or virtual platform) also takes shape to be exploited and disseminated as a basic concept of education. In fact the need to educate continues , because what was once widespread in relationships , incidentally, common : live contact and via forms (handouts / books) , today are also widespread in new relations, said virtual , supplanted by digital media - video conferences , handouts in PDF, etc..

This reflex is felt by all, whether in family relationships, whether from work, whether in school, and why not say the social experiences of the community (study groups, clubs, various associations, etc.). For discuss this from the perspective of information security in order to explore the digital citizenship is the purpose of this work. Let's face it:

2. Citizenship

Citizenship is a concept that has changed, as well as the definition of ethics, shaping - up over time given the standards of the people and its historical period.

The word derives from the Latin civitas (city), and formerly meant the political "status "of an individual in that community. Currently citizenship bias still gives the "status", but on the range of rights that the individual may enjoy and dispose. Initially binds to nationality (to belong to a country - individual who belongs to a Nation) , and citizenship as unfolding shows up in the use and enjoyment of political , relating to the rights of its powers the individual within that nation-state , comprising the basic rights and guarantees relating to the actions of the individual in his capacity as a citizen, as examples : vote , be voted , power meet in associations , participate in politics as a subject of law , enjoy the civil and social rights etc. .

In the broad sense, so has the concept of nationality, that like tell us the Professor Peter Lenza, it can be defined as *the legal and political bond that binds an individual to a particular state, making this individual becomes part of the people of that State and, therefore, enjoy law and subject to obligations.*

Since long before the spread of the internet the challenge of providing higher education to individuals reach these led efforts by governments and societies. A person with access to school and quality education is what you consider a position to exercise their rights, just as well educated

individual, therefore a person: a citizen in the fullness of the word.

There is this perspective formal education at school and seized accustomed to academic issues , culminating in the choice of a profession for later access to the labor market , the traditional education whose responsibility is the family or household , that of building character and sense of positioning of the individual between good and evil , right and wrong , can be anchored in religious doctrines, and there is a bias of this traditional education that honors the said " general knowledge", i.e., experiences or knowledge about different cultures , languages , food , clothing , character and history of a particular people, etc. .

In a broad aspect of citizenship dissemination of knowledge in the web of internet became more accessible for all included in the virtual world, getting governments to the challenge of providing such inclusion, which does not cease to be a major challenge for developing nations.

However, overcome this obstacle, with programs like " Access SP " of the State Government of São Paulo (to make the digital inclusion of needy population) , among many other prominent , as promoted by the Government of Uruguay - which provides Internet access in public places , or those most widespread in European nations - where access exists in all schools and public transport , it is understood that part of the problem is overcome.

A well informed and culturally educated citizens from their basic education tends to focus on range of situations, not just have shallow knowledge, this implies a greater knowledge about the everyday problems, improving its formal intelligence, emotional intelligence, his sense critic's making personal and professional decisions, among many other benefits.

The school (like insertion mechanism) and the teacher as content- not just professional, but mostly as an encourager of new discoveries together can act as a " differentiator " training of future generations of individuals , serving up virtual / digital methods .

The internet in our times is capable of spreading many cultural information and knowledge on this scenario is not inevitable conclusion that "digital education " is more of a tool for generating citizenship, because the internet has the ability to make relationships, although virtual, yet gain more depth and intensity, even approaching the citizen of his goal of knowledge, behold, there is spread with amazing speed.

For example, “Poupatempo” (Program of the State Government of SP causes the citizens nimbly get an identity document or a driver license, among others). In partnership with Access SP Program of digital inclusion for needy population, Citizens also feel the proximity of state presence and his work pro society.

In other sections you try , you can still think of monitoring for public safety and creating digital records of clinical hospital patients, so that any healthcare professional can, a large computer program to analyze data from an unconscious patient, example, it arrives au hospital after a hit.

However, the most important is to get to get to the citizen , especially in developing countries, basic education information and culture countries like know about rights and fundamental and human rights guarantees, so that it can be positioned facing problems experienced in community .

In Brazil there are many municipalities in a tool called "participatory budgeting”, mainly for communication between governments and poor communities, where individuals of that community may choose / decide where greater amounts of public money to improve collective environment that will apply. Example: they can decide between the priority of building a hospital, rather than a school.But in Brazil it is still not done completely by the Internet.

3. Conclusion

Finally, it should be clarified that the dissemination of information, especially in modern times where there is the prestige the virtual platform, it also owes much to the encouragement of education of the individual in seeking information, and not just wait for the government to provide learning situations.

Unfortunately this is still a terrible obstacle to be overcome by BRAZIL, where the vast range of its population, accustomed to the welfare self-indulgence of numerous governments, not easily predisposes to seek new learning mechanisms, especially on the internet, which unfortunately is still much more used to value social networks and less study.

4. References

[1] Bobbio, Norberto, “1909-A era dos Direitos”. Translation of Carlos Nelson Coutinho; apresentação de Celso Lafer.

Nova editora – Rio de Janeiro: Elsevier, 2004. – 13ª reimpressão.

[2] Pinheiro, Patricia Peck, Direito Digital. 4 ed. Ver. Atual. E ampl. São Paulo. Saraiva. 2010. p 43.

[3] Fiorllo, Celso Antonio Pacheco – Conte, Christiany Pegorari - Crimes no Meio Ambiente Digital – São Paulo: Saraiva, 2013.

[4] Cardon, Dominique, 1965 – A democracia Internet: promessas e limites. Tradução Nina Vicente e Tiago Coutinho; Revisão técnica Manoel Barros da Motta. – Rio de Janeiro: Forense Universitária, 2012.

[5] Bobbio, Norberto, "1909-A era dos Direitos". Translation of Carlos Nelson Coutinho; apresentação de Celso Lafer. Nova editora – Rio de Janeiro: Elsevier, 2004. – 13ª reimpressão.pp 15 e seguintes.

[6] Paesani. Liliana Minardi, Direito e Internet: LIBERDADE DE INFORMAÇÃO, PRIVACIDADE E RESPONSABILIDADE CIVIL. 5ed. São Paulo. Atlas. 2012.

[7] Piovesan, Flavia. Direitos humanos e justiça internacional: Um estudo comparativo dos sistemas regionais europeu, interamericano e africano – 4 e. revisada e ampliada. São Paulo: Saraiva, 2013, p 45.

[8] Sydow, Spencer Toth. Crimes informáticos e suas vítimas/ Spencer Toth Sydow; coordenadores Alice Bianchini, Ivan Luís Marques e Luiz Flávio Gomes – São Paulo: Saraiva, 2013 – (Coleção saberes monigráficos)p. 06/07

[9] Reale, Miguel. "1910 – Horizontes do direitor e da história". 3ª Edição, verificada e aumentada. São Paulo: Saraiva, 1999, pp.284, 292-293.

[10] Lenza, Pedro. Direito Constitucional Esquematizado. 17ª ed. Editora Saraiva. 2013

[11] Novelino, Marcelo. Da Cunha Jr, Dirley. Constituição Federal para Concursos. 3ª ed. Editora Jus Podivm. Salvador. 2012

KumuCloud: A Cyber Learning Tool

M. R. Zargham
Department of Computer Science
University of Dayton

M. R. Sayeh
Department of Electrical and Computer Engineering
Southern Illinois University

ABSTRACT
The cost of higher education is forcing some universities to restructure their traditional teaching practice and reduce cost by offering some courses online or in blended form. It appears that in future some universities may need to scale done in infrastructure by increasing the use of technology in their course delivery. It may become to a point that first year students are required to come to campus most of the time, however as they progressing toward their senior year, they may require to be on campus lesser and lesser every year by utilizing more online materials. One of the free sources for them is open access educational materials. We have developed an e-learning portal, called KumuCloud, (Kumu in Hawaiian language means "Teacher") to address these needs. This e-Learning Portal web site "www.KumuCloud.org" provides learning modules to everyone free of charge through YouTube and many other sites. It helps the students to get high quality education material very quickly at anywhere and anytime.

Keywords
Open Access Education, Open Educational Resources, Mobile learning, e-learning, Educational Site, Teachers in Cloud, KumuCloud

1. INTRODUCTION
Open access education (or open educational resources) was first introduced in UNESCO's 2002 Forum to enhance the higher education in Developing Countries [10]. The idea was to provide learning modules to everyone free of charge under an open source license. The license allows the users to reuse, revise, remix, or redistribute such learning modules. Today, there are many sites to provide free lecture videos. Many of these free lecture videos are from different universities or open access educational sites. Sooner or later majority of students will tap into these resources to reduce their cost while increasing time and location flexibility.

However many of these sites have either limited resources for course development or are very specific toward one way of course delivery. To address these deficiencies, our e-learning portal site, KumuCloud.org, provides a belended environment for learning. In addition to top lectures from around the world, instructors can record their own lectures and make it available to students within an hour from the lecture time. With placing these videos on KumuCloud, students can review the lectures through their smart phones in addition to their computing devices. This access will help in situations where some students may have confusion regarding a topic after the class and while the professor is not available; some might miss a lecture and want to have access to the corresponding lessons or some may have an interest of an entirely different topic which they have never studied and would be able to access it via KumuCloud. They can also communicate with each other to discuss the lectures or project requirements.

2. KUMUCLOUD COMPONENTS
This section describes the main pages of the KumuCloud. Figure 1 represents the home page for the site. The main component of this page is explained below.

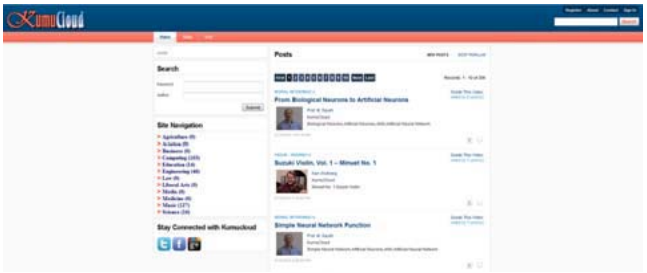


Figure 1. Home Page of KumuCloud

Post Category: Currently the portal focuses on three types of posts. They are: a) Video, b) Slide (e.g. PowerPoint presentation) and c) Text. These categories make it easier for the filtering logic so that students can go directly to the targeted category and search the post they want.
Post Navigation: To get more precise posts on a specific area, site navigation can be very useful. It supports a four step navigation system.
Grade of the Post: Grade represents the quality of the post. It is very important since it separates high-quality posts from poor-quality ones. So students can easily figure out which are the useful posts among thousands of posts. A novel ranking algorithm has been used to sort the articles.

Once a user has selected a particular department, a page will appear listing related topics/courses. The user can then select a particular topic within the possible choices. Figure 2 represents a sample page for a topic in the field of computing. The main features of this page are explained below.

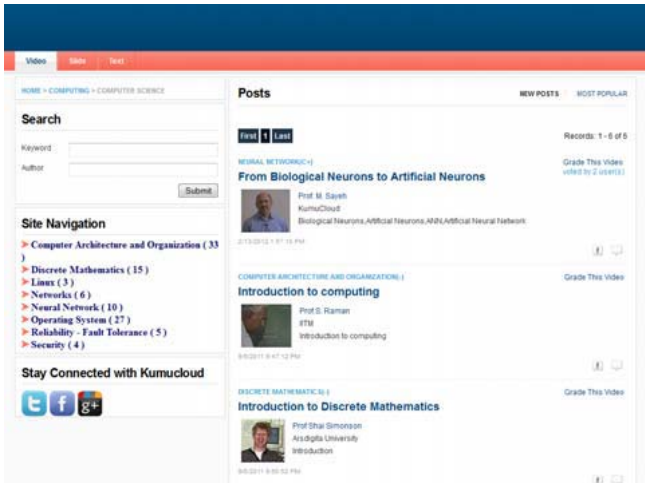


Figure 2. Listing of KumuCloud

- *Breadcrumb*: Helps the user to navigate back to the previous categories. The above figure is showing the student is looking the posts which are related to the "Data structure" course which is one of the courses of the "Computer Science" department. Clicking on "Computer" will redirect to the previous navigation and will show the list of all courses in the Computer Science department.
- *Keywords/Areas*: There are many sections/areas/topics of each course. As soon as the user selects the course from the navigation, it displays the list of topics of the course which will filter the posts to get to the right post quickly.
- *List of the Posts*: It displays the list of the posts based on the user's selections.
- *Share in the Community*: Furthermore, the user can share a high-quality post in their own community (e.g. Facebook, Twitter, Google+, etc.) to notify others about it. It helps to spread the existence of new materials among people quickly.

When a video (or slide, text) is selected, a page representing the selected video will be opened. Figure 3 represents a sample for this page. The main component of this page is explained below.

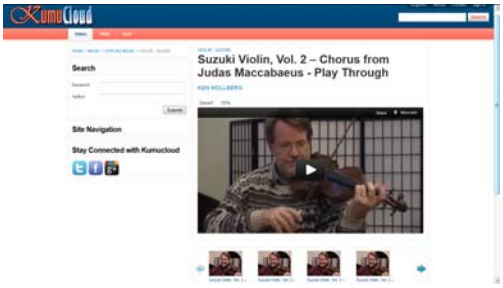


Figure 3. Detail Page of the Post

- *Area/ Keyword*: Displays the keyword/area of the post which gives some idea of the purpose for the post to the user.
- *Title of the Post*: It displays the title of the post.
- *Main Content Area*: This is the description area. In the case of a video/presentation post, it displays a video player which streams the data to the user. In the case of a text post, this area contains the descriptive text of the post.

Figure 3 represents a music video (playing Chorus from Judas Maccabaeus) which can be played in 100% and also 50% of original sound. By selecting 50% on the top (left corner) of the image, the video will be played in half speed, otherwise will be played in full speed. There are two types of videos: tutorials and play through. Once students have learned bowing and correct notes from tutorial videos, they can utilize play through videos at different speeds for exercise.

As mentioned above, KumuCloud have three different categories of posts: video, slide and text. Each of them can be linked with others. Figure 4 shows a window where a slide is linked with a video and text. If a post is linked with other posts then it displays different colors of bubbles here. Each color of bubbles has different meaning. A red bubble used for video. Similarly a green and blue bubble is used for text and slide, respectively. By clicking on the bubble, user can navigate to the linked post. In case of slide and text post, this file link section will display the link of the file. User can view/download the file by clicking on the link.

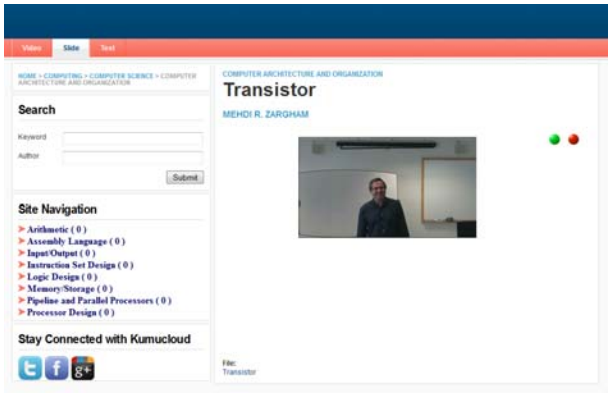


Figure 4. Linking of different posts.

3. A BLENDED COURSE EXAMPLE

In this section, we will layout the plan for teaching an introductory computer science course in a blended environment. The students will be involved in various training and social activities throughout the course. Students will be encouraged to develop their projects on their phones.

During the first month, an Android smart phone with Wi-Fi capability will be loaned to each team of students for the rest of the semester.

During the first four weeks of the class, we plan to have students involved into a group project, each group is made of 2 to 3 students, mixing of strong and weak backgrounds students and making every effort to have one minority student in each team when is possible. This prevents strongly motivated students become bored and at the same time provides more opportunity for weak students to succeed. We will regularly monitor their progress and make sure that they do task sharing and collaboration. We meet with students once a week for at least two hours. Students receive all the training through the KumuCloud. Our weekly meeting will focus on answering questions and going through main subjects. During the first four weeks, in order for students **become familiar with app development on their smart phones**, we offer them a template for creating some simple projects. One of simple project for students could be calculating body mass index for any individual. This app and couple of other with more complexity are explained below. Students develop these apps in Eclipse environment and simulate by Android. They, then, upload these apps on their smartphones or share them with community. Students can instantly see their product to be used or shared with the world. With these motivation and encouragement, they are motivated to learn further about computer science. We also allow some teams based on their background and interest use other software development platforms such as Matlab. These students may be majoring in Engineering or Physical Sciences. They will be working very close with the mentors (Upper level or graduate students) who are very familiar with the chosen platform.

Examples of Projects

Body Mass Index (BMI) **(all teams).** This app accepts the weight (in ponds) and height (in feet) from users and displays their BMI. It first converts the entered weight and height to kilograms and meters, respectively. Then calculates BMI by dividing weight by height squared (i.e. $BMI=Kg/m^2$). Based on the BMI value, the app displays one of the appropriate below terms for the user.

Adults	Women	Men
Anorexia	<17.5	
Underweight	<19.1	<20.7
In Normal Range	19.1-25.8	20.7-26.4
Marginally Overweight	25.8-27.3	26.4-27.8
Overweight	27.3-32.3	27.8-31.1
Very Overweight or Obese	>32.3	>31.1
Severely Obese	35-40	
Morbidly Obese	40-50	
Super Obese	50-60	

To do this project, students are taught through a detailed list how to install development tools for Android. All the tools, such as JDK, Eclipse, ADT, SDK, and Google API, can be downloaded from Internet for free.

Campus Tour **(two teams to three teams).** In every college, there is a new student admission office that often provides campus tour to prospective students and their parents. The goal of this project is to do this task by developing a self-tour guided app that displays a list of the five nearest building using GPS technology. The distance to each building will be displayed as well as each building name. A video about each building is placed next to the name in case the user wishes to see which building the name is referring to and also to know the building's history and the units it includes. The app also shows the current position of the visitor at any moment on the campus map.

Given the fact that implementation of this project is hard for many beginners, we provide students with a basic sample app and ask each team to improve the code.

Survivor game. There are two types of Orbs, blue and red. There are ten of each kind. Each kind of Orbs needs to hit another Orb in order to survive. The attacked Orb will die. An Orb can hit another Orb by hitting its sharp edge to the round part of the other Orb. The stronger Orbs will survive!

Each Orb is a neural net. They are capable of self-improvement in order to become stronger. Each Orb can see the location and direction of the closest Orb of the different type. Each Orb is able to do four actions, move forward, turn left, turn right, and hit.

Vacuum Robot.In **t**his project you will program a simulated robot to navigate through its environment. The robot senses its environment by using bump sensors (in case of hitting objects or walls), global position, dirt sensors. With the data from sensors your program should decide an appropriate move for the robot to execute.

Maze Robot. In this project you will program a robot to find its way through a maze going from a starting point to an end point. A maze is a grid of square cells where each cell is a wall or an empty path. The robot should travel through empty cells to reach to its destination. At each state of the movement, the robot can see (sense) its four neighboring cells (South, East, North, and West) and based on these cells it should decide its move direction.

Gender ClassificationThis project classifies male and females based on their full-body pictures. Studentswill address the following specific tasks: 1) Constructing a data set by asking many volunteers to have their pictures taken from a front angle, full-body shots. They will take full-body shots from a specific distance with the subject standing completely erect against a white wall. White background area is used to simplify feature extraction in later stages; 2) Learning MATLAB; 3) Selecting important features (such as hair, shoulders, height, body mass, etc.) and extracting them from images by using MATLAB. The major difficulty here will be to eliminate noise as much as possible and reduce the number of objects in the picture in order to have more accurate measurements of selected features of a given person. 4) Once they have decided which features to use and how to extract them, their next step will be to train a model to distinguish gender.

Zoo Tour.In Cincinnati, Ohio there is a beautiful Zoo that attract many tourist and local people. Unfortunately, there is nobody to offer a comprehensive tour to visitors. The goal of the first phase of this project is to do this task by developing a self-tour guided app that displays a list of the five nearest cages using GPS technology. The distance to each cage will be displayed as well as each cage name. A video about each cage is placed next to the name in case the user wishes to see which cage the name is referring to and also to know detail information about the animals that cage includes. The app also shows the current position of the visitor at any moment on the Zoo map. However, in practice, it is possible that the Zoo replace the animals in a cage without updating the app's GPS coordinates. To address this problem, in the second phase of this project (perhaps in another course), the students modify the app by including machine learning capability. If the user has doubt about the residents of a cage, he/she may take a picture the resident animal in order to be recognized by the app.

4. CONCLUSION

In this paper, we have described the main components of an open access educational site, called KumuCloud. We are hoping that in near future, this site become a major source for any college student around the world to access course lecturer videos/ text/ slides on many different fields of study. Considering the progress in e-learning throughout the world, this site may function as a great supplement, especially for underrepresented group. Among the differences between this site and the existing ones, there are three main differences that make the site more practical; these are 1) the materials are reviewed before they are published, 2) The materials are ranked based on the user ratings through a novel method, and 3) The materials are not limited to some specific sources rather they are compiled from around the world.

5. REFERENCES

[1] Journal Knight, S. 2005. "Innovative practice with e-Learning. Bristol", JISC e-Learning Programmed.

[2] Web Site teaching and learning. http://home.sprynet.com/~gkearsley/engage.htm

[3] Web Site http://en.wikipedia.org/wiki/Bayes%27_theorem

[4] Web Site http://leedumond.com/blog/the-wisdom-of-crowds-implementing-a-smart-ranking-algorithm/

[5] Chandra, S., 2011. "Experiences in personal lecture video capture", IEEE Trans. On Learning Technologies, Vol. 4, No. 3, pp. 261-274.

[6] Sharples M. and J. Roschelle, 2010. "Guest Editorial: Special issue on mobile and ubiquitous technologies for learning", IEEE Trans. On Learning Technologies, Vol. 3, No. 1, pp. 4-5.

[7] Kukulska-Hulme, A., M. Sharples, M. Milard, I. Amedillo-Sanchez, and G. Vavoula, 2009. "Innovation in mobile learning: a European perspective," Int'l J. Mobile and Blended Learning, Vol. 1, No. 1, pp. 13-35.

[8] Leinonen, T., J. Purma, H. Poldoja, and T. Toikkanen, 2010. "Information architecture and design solutions scaffolding authoring of open educational resources," IEEE Trans. On Learning Technologies, Vol. 3, No. 2, pp. 116-128..

[9] Thoms, B., 2011. "A dynamic social feedback system to support learning and social interaction in higher education," IEEE Trans. On Learning Technologies, Vol. 4, No. 4, pp. 340-352.

[10] Duval, E., and D. Wiley, 2010. "Guest Editorial: open educational resources," IEEE Trans. On Learning Technologies, Vol. 3, No. 2, pp. 83-84.

Teaching IPAD Forensics Online With a MOOC

Professor and Cybercrime Training Lab Director Eamon P. Doherty CCE, CPP, SSCP
[1]School of Administrative Science, Fairleigh Dickinson University, Teaneck, N.J., USA
Doherty@fdu.edu, FDU, 1000 River Road, H-DH2-16, Teaneck, NJ, 07666

Conference : EEE2014

Abstract **-** In 2013, approximately one hundred and twenty two million people have an IPAD. Many of these devices will hold evidence of a white collar crime or hold data about a victim of a crime. These new devices will need to be examined. The paper discusses the need for a MOOC that teaches someone how to do IPAD forensics and then allows them to print a certificate. The content includes instruction on what some people consider the "best practices" in IPAD forensics along with a brief survey of some commercially available and open source tools used by practitioners. Some options for testing, the issuing of certificates, and performing outcomes assessment are also explored in this paper.

Keywords: IPAD, IPAD Forensics, Teaching e-Forensics, Online Learning

1 Introduction

It is very important to have online classes since many National Guardsmen, policemen, and security contractors have ever changing schedules [1]. Some students who are National Guardsmen or defense contractors often discuss anecdotes of being deployed on short notice to areas where living conditions are basic and commuting to traditional classrooms is not possible. These students often express a sentiment that they are glad that online classes exist. Since the summer of 2013, some digital forensic examiners on the Certified Computer Examiner (CCE) Listserver have expressed a need to learn what tools and methods can be used to examine an IPAD. Online learning is one low cost option for them. A MOOG may be a good environment for teaching IPAD forensics.

A study quoted by Amy Gesenhues said that are one hundred and twenty two million IPAD users as of 2013 [2]. Michael O'Grady has a report that says that 60% of North American Consumers will have an IPAD by 2017 [3]. Since many professionals, students, and consumers have such a device for activities of daily living, then it is only natural that a subset of this one hundred twenty-two million IPADs would be a victim or perpetrator of white collar crimes. IPADs also synchronize with a variety of computing devices such as notebooks, laptops, and desktop computers. There may be occasions to examine all of a person's devices in order to find exculpatory evidence to prove a person is innocent. It should become more apparent that IPAD forensics is becoming more necessary and that training needs to available online for those investigators with immediate needs or limited budgets for travel and traditional classroom training.

2 Challenges Using Traditional Forensic Tools

The IPAD uses an Apple IOS operating system and not the more common Microsoft operating system that uses a FAT 32 file system. This departure from the usual file systems and family of operating systems means that digital forensic examiners will need to learn about the new platform, acquire new investigation tools, and learn to use them. This presents an opportunity for educators to provide a service to investigators by creating a low cost class with learning materials that include links to a digital image of an IPAD, with simulated criminal data, that they can practice examining. The class should also contain links to tools that are accepted by the computer forensic community for use with examining an IPAD. Since the IPAD is relatively new and some of the tools for examining an IPAD are also new, it would behoove the examiner to learn what tools have been successfully used in court cases. Tools and methodologies accepted by the scientific, academic, legal, and law enforcement communities are said to have passed the FRYE test. "Citing Frye v. United States, 54 App. D.C. 46, 47, 293, F. 1013, 1014, (1923), the court stated that expert opinion based on scientific technique is inadmissible unless the technique is generally accepted as reliable in the relevant scientific community [4]." Tools and methodologies that have passed the FRYE test are more likely to be accepted in future cases thus allowing the examiner to focus on the case at hand, and not be preoccupied with the tools.

It is also important to know that all IPADs are not the same. There are many versions of the IPADs and not all tools work with each version. It is important to make sure that if you have an IPAD 1, IPAD2, IPAD 3, IPAD 4, or IPAD Mini, and that your tool works with that version of the IPAD. The examiner should also be familiar with the dates for each generation or he or she could be made a fool of in the

courtroom during the Voir Dire process when a person is trying to be accpeted as an expert by the court. If there is a discussion of the history of the IPAD, it would behoove the examiner to look at similar Apple devices such as the Apple Newton MessagePad 2100 and the Apple Graphics Pad from 1979. A discussion of the history of a product could definately arise in the courtroom when there is a discussion of the examination tools and the device in question. It is better to be overprepared than under prepared. A person who is not aware of the history of a device is taking a risk of appearing to be unprepared and not deemed an expert.

3 Possible Tools For IPAD Forensics

Paraben is a company that has created and sells a digital forensic tool called Device Seizure 6.0. This tool can be used for collecting, analyzing, and reporting on evidence from a variety of smartphones, PDAs, Garmin and Tom Tom GPS Navigation Devices, and IPADs [5]. There is also an opportunity to get a demo model of the software too. Since Paraben has a number of products that have been successfully used in digital forensic examinations, this would certainly be a good candidate for use in the classroom and in the field. There is also telephone and email support for the tool which makes it an attractive option for both investigators and educators.

Let's consider something different, the free open source tool. Many people like open source tools because they can customize them. Having the source code for a tool allows the investigator to learn exactly how the tool works. This is very important if the investigator is asked in court about how the tools work. There are usually online communities of people in the open source tools community who can help each other create new functions or provide each other with technical support. If the educator and students have a computer science background, then an open source tool may be a good option for both teaching and usage in the field. Johnathan Zdziarski released a new open source tool called "Waterboarding" which is made for examining ios devices such as the IPAD. An article says this about Waterboard, "Waterboard is an open source iOS forensic imaging tool, capable of performing an advanced logical acquisition of iOS devices by utilizing extended services and back doors in Apple's built-in lockdown services. These service can bypass Apple's mobile backup encryption and other encryption to deliver a clear text copy of much of the file system to any machine that can or has previously paired with the device. Acquisition can be performed via usb, or across any wireless network where the device can be reached [6]."

Lantern 3 is an ios forensic tool that can be purchased from Katana Forensics. It is another investigation tool option for both educators and investigators. Lantern 3 is good for investigating IPADs, IPhones, and other ios devices [7]. There is a demo version of the software that can be used for learning purposes. If the student is serious about the subject and wants to engage in IPAD forensics as a career or sideline, then an investment in the full priced product can be made by the student. The interface is not a command line interface but rather a graphical user interface with pull down menus. The pointing, clicking, and built in documentation can make it more intuitive for people with a limited computer background.

Let us consider another commercial tool made specifically made for the security and investigation community. Oxygen Forensic Suite 2012 (OFS 12) is a tool that law enforcement frequently uses for both IPhone and IPAD investigations. Oxygen Forensics does have a limited time version of their software that could be useful for both educators and students for either an online or in person class. The OFS 12 can be used for the IPAD 2 and the new IPAD [8]. The tool has a variety of analytics that can be used for creating a timeline and analyzing both the incoming and outgoing phone call activity of people. If students are interested in learning more about this tool, they can contact the company and find out what trade shows that the vendor will attend. There are often times that students can also see demonstrations of a variety of mobile device forensic tools at conferences such the High Tech Crimes Investigative Conference (HTCIA) or the Mobile Forensics World.

Black Bag Technologies has a tool called Blacklight Forensic Analysis Software. This is very useful for security professionals who need to preserve evidence, collect it, analyze metadata, and then make a report. There are filters that allow one to collect all the pictures with GPS metadata embedded in them. Being able to narrow data down to file type is important in cases where a person with a top secret clearance might have been accused of viewing child pornography. The filter would let the investigator find only the images which could then be searched for the contraband while leaving all other top secret documents alone. There is also an option on the IPAD to wipe all data if the password is incorrectly guessed ten times. This is reminiscent of the Blackberry phones. It may be necessary to seek a compel order from a judge in order to retrieve the password from the suspect.

4 Motivating Students

There are occasional times when students will ask about possible jobs for mobile device forensics. Some students appear quite mercenary and only engage their full efforts in classes and majors that have a high probability of earning a high wage. Showing requests for IPAD and IPhone forensic experts from the CCE Listserver appears to allay such fears about investing themselves in this area of study.

5 The Examination Machine

Some forensic examiners who are practitioners have mentioned in continuing education classes that it is best that the examination machine is wiped with a pattern of ones and zeroes after each examination. Then a new forensic image of the examination machine with its standard tools and operating system is reloaded after each examination. This insures that no rootkits or zero day exploits are on the examination machine.

It is also important that the computer that is used to examine the IPAD is free from viruses, spyware, and malware. It would behoove the examiner to run a copy of antivirus software and antispyware software on the examination machine before using. This way the computer is free of root kits, data diddlers, spyware, and other malware that compromise the examination by changing evidence or what the examiner sees on the screen. The examination machine should be taken be offline and not be accessible to hackers by WiFi, Bluetooth, wire, infrared, or any other means.

It is also important to tell students that all the software on the examination machine is currently licensed to the examiner and that everything is up to date. There should be no pirated software because that could nullify the results of the examination with a legal concept known in the United States as "Fruit of the Poisonous Tree." The list of programs, serial numbers, and license information should be in a spreadsheet and ready for display to both sides in court.

6 Isolating The Evidence

Before we isolate our evidence we must first identify it. Did we collect any type of storage devices or media that associate with the IPAD? We may wish to bring a wireless device as well as some signal detection equipment to the crime scene to see if there are any hidden storage devices hidden in a false ceiling or cardboard boxes. We may only get one opportunity to visit the crime scene and take all the evidence.

If the IPAD is seized and transported to a facility for examination, it is also important to isolate it from other networks that will change date and time stamps of the device and bring doubt about the integrity of the evidence. A faraday bag should be used so that the device is isolated and not influenced by outside signals. It is very good to have lessons that reinforce this concept of keeping evidence tamperproof.

7 Permissions

It is very important to reinforce the need for permission to examine an IPAD. Consent could be given by a person. In a law enforcement investigation, a search warrant would be in order. In corporate investigations, it is important to check that the device belongs to the corporation and that the person using it has no expectation of privacy. Signed policies filed with human resources and advisements from general counsel are highly desirable in corporate investigations. It is also good to advise students about the American Fourth Amendment Exception which allows the United States Customs and Border Patrol to search a person's IPAD when leaving or exiting the country [9,10]. It is important to tell students that this is not some type of modern loss of freedom but something that goes back to the beginning of the nation.

7.5 Chain of Custody

It is also important that the class materials contain information about the chain of custody which is also called a chain of evidence. The chain of custody contains the history of the evidence from the time that was seized until the time that it appeared in court [11]. It is also important to teach about preserving evidence and making sure that it was not spoliated due to carelessness or ignorance. The chain of custody should show the audit trail of a piece of evidence and must contain the case number, the examiner's name, and other information such as an evidence description. Students need to understand the importance of the chain of evidence and that it is one of the first things that defense lawyers target since it is an easy win for them if someone was sloppy.

7.7 Navigating the Deposition

Student materials should include what is a deposition and how to answer questions during a deposition. Videos such as this one show trial attorney experts who discuss the importance of answering questions correctly and what to avoid.
http://www.youtube.com/watch?v=6Y9uNojuwjw Students who will do IPAD examinations as part of their profession will most likely have to give a deposition about their findings, how they obtained those results, and what methods they used. Depositions can be grueling and students should have some preparation as well as know places where they can seek additional training. Some training will discuss how to dress while other videos may discuss controlling body language and how to answer questions. All are important but a diversity of videos is advisable http://www.youtube.com/watch?v=PuYVhIa_vIg

8. Creating a Test Plan

It is also important to have a test plan for examining the IPAD. The test plan should include the methodology for seizing the device, collecting the evidence, analyzing the evidence, and reporting on it. The test plan needs to have pictures, screen shots, and rationale for doing each major step in the investigation. The rationale should also discuss the selection of certain tools and how they may pass the Frye Test. The class materials should include this North Carolina Court video of a cell phone forensic expert who did not have a test plan to show the court. The students should be encouraged to watch from the first twenty minutes of the film. The test plan discussion starts at 10 minutes and 30 seconds and can be followed for a significant period of time. It becomes apparant that such a document as a test plan is important for helping to maintain one's credibility and good standing in court.. http://www.wral.com/specialreports/nancycooper/video/9494914/#/vid9494914

9. Validating Your Forensic Tools

Another part of the test plan should include verifying the forensic tools. Amber Schroader wrote a booklet called, "How to Validate Your Forensic Tools." This booklet is not for sale but can be obtained free by contacting the Paraben Corporation. One of the questions that the booklet asks is, "Is the forensic technology read only?" If the answer is yes, then the examiner has less worries about changing evidence. If the answer is no, then perhaps an option such as using a USB Forensic Bridge Writeblocker might need to be added so the answer becomes yes and evidence is preserved. Amber asks if we can repeat our results. Anyone who studies science has heard that the scientific method requires that our results be repeatable. The same is true in digital forensics. Can the defense or prosecution take our forensic image and tools and get the same results? Another basic question from Amber's book is," Was the tool was designed for forensics and are the images valid? This is a good question because an examiner should not use a tool that might have been posted by hackers and contains malware or malicious code.

10. The Structure of the Class

A continuing education department could advertise the IPAD forensics class and then charge fifty dollars. The people could sign up with a credit card using a continuing education website and shopping cart. A username and password for Blackboard could be assigned to the person who paid. The person could log in and read a pdf file with all the class materials. He or she could watch some videos. The next step would be to access links for Lantern3, Waterboard, Device Seizure 3, and the Oxygen Forensics software. The student could fill out the information and obtain the limited time and limited functionality software. A forensic image of an IPAD with simulated criminal data could be provided by the instructor and then be downloaded. The student could use all the tools with the forensic image and find all the data, analyze it, and make a report. Then the student could take a test. If the student obtained a grade of eighty percent, then a passing grade could be obtained and a certificate with his or her name could be printed. A limited number of tries at the test are recommended so that the person does not repeatedly try combinations of answers until he or she passes. Unrestricted testing could severely limit the credibility of the class and its alumni if such testing practice were uncovered during the voir dire process in court. Voir dire is the process by which a person and their experience, education, and knowledge are securitized in court for the purpose of determining if he or she is an expert witness [12]. It is also good to refer students to a short book by Stanley Brodsky that discusses how to dress for court, how to testify, as well as the emotional / psychological aspects of testifying in court [13].

If one does not wish to provide a forensic image of an IPAD for the student to download, then the student could just examine his or her own IPAD with the forensic tool links provided. Then the test would be about using the tools, forensic procedures, and best practices. This may be a better option because it would focus on general knowledge and not just be a treasure hunt for data on the forensic image of the IPAD.

11. Outcomes Assessment

It would also be useful to perform an outcomes assessment process with this form of learning. The course would need a set of learning objectives, recommended literature for further study, and a list of the vocabulary that is relevant to this study that is included in the course. It would also be useful to do a quiz at the beginning of the MOOC and the same quiz at the end to assess an increase in knowledge. In order to improve the class, there should be a digital mailbox for collecting suggestions on how to improve the class.

12. Conclusion

Teaching students to become a forensic expert about the IPAD requires academic knowledge and social knowledge. The Stanley Brodsky book teaches the social aspect while the Rebecca Bace / Fred Smith book teaches the academic and legal concepts necessary for being an expert in court. It also important to teach students how to preserve, identify, analyze, and report on digital evidence. Students also need to know how to use commonly used tools such as Lantern3, Waterboard, Device Seizure 3, and Blacklight Forensic Analysis Software. Lastly, the chain of custody, how to give

a deposition, and the basics of the operating system and file systems should be explored. Multimedia, trial versions of software, and online learning environments make learning possible for those who need to learn about topics such as IPAD forensics.

13. References

[1]. Doherty, E. P. (2011), Teaching Cell Phone Forensics and E-Learning, 2011 International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE'11: July 18-21, 2011 Las Vegas, Nevada, USA). Conference Proceedings, Pages 161-5, ISBN 1-60132-176-7

[2]. URL Accessed 10/17/2013 Gesenhues, A., (2013)," Study: Android Tablets Surpass iPads In Q1, Tablet Usage Up 282% Since 2011", http://marketingland.com/study-android-tablet-users-surpass-ipad-users-by-34m-in-q1-with-tablet-usage-up-282-since-2011-2-47740

[3]. URL Accessed 10/22/2013 O'Grady, M., "Global Business And Consumer Tablet Forecast Update, 2013 To 2017" **http://blogs.forrester.com/jp_gownder/13-08-02-global_business_and_consumer_tablet_forecast_update_2013_to_2017_0**

[4]. Smith, F. C., Bace, R. G., (2003), "A Guide to Forensic Testimony, The Art and Practice of Presenting Testimony as an Expert Technical Witness", Boston, MA: Pearson Education Inc., Addison Wesley, ISBN 0-201-75279-4, p 407

[5]. URL Accessed 10/22/2013 http://www.paraben.com/downloads/ds6-supported-devices.pdf

[6]. URL Accessed 10/17/2013, iClarified, "Jonathan Zdziarski Releases Waterboard: An Open Source Forensic Acquisition Tool for iOS Devices http://www.iclarified.com/31060/jonathan-zdziarski-releases-waterboard-an-open-source-forensic-acquisition-tool-for-ios-devices"

[7].URL Accessed 10/17/2013

http://katanaforensics.com/lantern

[8]. URL Accessed 10/20/2013 http://www.prweb.com/releases/iphone5-ios6/mobile-forensics/prweb10067611.htm

[9]. Wilson, V.,(2011)," Laptops and the Border Search Exception to the Fourth Amendment: Protecting the United States Borders From Bombs, Drugs, and the Pictures from Your Vacation",

[10]. University of Miami Law Review, pp 999-1002, Vol. 65:999, 27-May-2011 URL Accessed 10/29/2013 http://lawreview.law.miami.edu/wp-content/uploads/2011/12/v65_i3_vwilson.pdf

[11]. Easttom, C., (2014),"System Forensics, Investigation, and Response", Jones and Bartlett Learning, Burlington, MA. pp 84-85, ISBN 978-1-284-03105-8

[12]. Bace, R., Smith, F., (2002), "A Guide to Forensic Testimony: The Art and Practice of Presenting Testimony As An Expert Technical Witness", ISBN 978-0201752793, pp 8,9,276

[13]. Brodsky, S. (1991), "Testifying in Court: Guidelines and Maxims for the Expert Witness ",Published by the American Psychological Association, ISBN-13: 978-1557981288

In-Depth Assessment with Multiple Choice: The Reverse Multiple-choice Method

Indu M. Anand
Sushila Publications
Chelmsford, Massachusetts, USA

Abstract -- *An assessment methodology based on "Reverse Multiple-Choice Method," is described. The novel method combines the ease, potential for automation, uniformity, objectivity and scalability of familiar multiple-choice format with the depth of feedback generally associated with essay or open format testing. The method is flexible and general, and may be adapted for educational testing in most subject areas at achievement levels from kindergarten through college.*

Keywords -- testing, multiple-choice, training, in-depth assessment

1. Introduction

Frederick J. Kelley invented multiple-choice tests exactly 100 years ago in 1914. In these one hundred years, multiple-choice testing method has been exulted and maligned, but steadily grown to become the world's most common method of educational testing, and the norm for examinations taken by large numbers of students. This run-away use of multiple-choice is beyond the inventor's original intent. The format was devised to address the then need to deal with an eight-fold increase in the number of college bound high school students within two decades. The numbers, no doubt are behind its dominance to this day, along with the convenience, low cost, high speed, and uniformity with which these tests can be administered and graded.

But, as we show here, the format has other non-obvious virtues not commonly appreciated. In fact, despite the oxymoronic juxtaposition of "multiple-choice" and "in-depth" in the paper's title, multiple-choice paradigm *can* be used for in-depth learning and assessment, by a *change of perspective*.

By way of a warning, Reverse Multiple-choice Method presented here, is *not* an *easy* alternative to traditional multiple-choice. It is not suitable to conduct an entire test such as the SAT. But it can be effectively used in several situations, including for instance, an instructor looking for *reliable* feedback from the students, or for an assessment imperative when mastery of the material is crucial, e.g., for the teacher trainees. For the SAT's and ACT's a few RMCM questions could be added to provide "control" of validity, something the "essay" portion was intended to do.

2. Multiple-Choice Format and Alternatives

2.1 Multiple-choice

A typical multiple-choice question comprises three sections: a set of presumed facts (e.g., a narrative, an expression, an equation, a geometric figure), an interrogative sentence (a.k.a., the "call of the question"), and answer choices, typically between three and five, one which would be graded as "correct" or "best" answer.

2.2 Open Format, Long/Short Essay

The classical methods of testing require a student to demonstrate in his own words mastery of the material learnt, by writing an essay on a given topic, by detailed answers to pointed questions or by giving solutions to problems given in a test etc.

In order to keep the students from rambling and keep answers within reasonable lengths, limits are often placed on the length of an answer.

2.3 Other formats

Several other formats have been devised and employed, for instance "filling in the blanks,"

"verifying statements," often as true or false, "matching answers to similar or related questions" etc. In some cases, a multiple-choice question may admit more than one correct answer, in other cases "all of the above" or "none of the above" are provided as possible answers, thereby discouraging random selection of answers. These methods are generally either variations of multiple-choice or attempts to replicate the ease and scalability of multiple-choice tests. They may also aim for automatic grading since computers are widely available for the purpose. We do not discuss these alternatives further but our analysis applies with modifications.

3. Advantages and Drawbacks of Multiple-Choice

3.1 Advantages for Examiners, and for Examinees

Multiple-choice tests are scalable and "objective," in that the grader may not contaminate grading and mark an answer correct or incorrect based on subjective bias; other advantages, for instance, uniformity, automatic grading, or low cost flow as byproducts. In fact, "democratization" by "objective" tests was recognized early as a virtue of the format.

The format has advantages for the test maker as well as the test-taker. For an examinee it means no tired fingers writing long answers and no time wasted in making the answer presentable stylistically or in penmanship.

For the examiner the advantages are immense: the drudgery of reading a stream of similar answers is reduced, even eliminated by machine grading. Due to the brevity of answers, a larger swath of the subject matter can be tested, leading to wider-scaled feedback. Pressure on test-takers is reduced as well, since spotting a correct answer is often easier than generating one.

3.2 Disadvantages of Multiple-choice and Its Alternatives

Advantages of multiple-choice format mentioned above come at a cost and lead to its drawbacks. The format provides little room for demonstrating the soundness of one's underlying analysis or the accuracy of computation. Total loss of credit may result from incorrect reading of a single word, phrase, value or fact in the question on which the answers turn. The *very* ease and simplicity of selecting and recording an answer to a test question may blur the difference between knowledge and ignorance.

Additionally, the format can encourage students to learn the material superficially. And worse, the examiner cannot easily spot cheating when it occurs.

On the other hand, while providing opportunity to test-takers to demonstrate their command of the subject matter, open format testing is time consuming to take and to grade, as well as expensive, subjective and non-uniform in measurement. Several other alternatives listed above suffer from similar disadvantages.

3.3 Reverse Multiple-choice Method: A Promising Alternative

The Reverse Multiple-choice Method (RMCM) offers a promising assessment alternative that combines the uniformity, efficiency and grading ease of "objective" or "standardized" multiple-choice tests with reliable "measure of knowledge and understanding" generally associated with open format tests.

RMCM assessment methodology lends itself to automatic grading which makes it attractive for distance learning and MOOCs, massively large online courses.

4. The Reverse Multiple-Choice Method (RMCM)

4.1 How RMCM Works

The Reverse Multiple-Choice Method utilizes multiple-choice questions, with a twist: It *requires* a student/examinee to consider each answer choice, and inquire how the given *facts* would be modified to make that answer choice the "correct" or "best" answer.

The distinguishing steps typically undertaken by an RMCM system are as follows:

- prompt the examinee to select an answer choice as the correct answer;
- record the examinee's selection and assign credit for the selection;
- prompt the examinee to select at least one of answer choices not selected as correct, then prompt the examinee to provide a follow-up query to which this selected answer choice is a correct answer;

- match the follow up query against stored queries for which the answer is correct and provide a score for the question depending on the match.
- The system can flag the answer for human evaluation when the follow up query does not match a stored query – so as to address a situation where an examinee's follow up query may be unanticipated but appropriate.

4.2 An Example

Let us consider as an example one of Frederick Kelly's original questions:

Q. Which of following animals is a farm animal?

(A) a cow. (B) a tiger. (C) a rat. (D) a wolf.

The correct answer is (A), which the standard multiple-choice question seeks. Here, RMCM question would further ask the student to determine how the query would have to be changed so that the correct answer would be (B), (C) or (D).

All answers here are animals, so the answer turns on the qualifier "farm." Therefore, changing the word "farm" is necessary, which requires one to think of the unique properties of cow, tiger, rat and wolf, not shared by other animals named. A rat is small, for example, unlike the other three animals. Tiger and wolf are both wild animals, but they can be distinguished possibly as being feline and canine respectively.

Therefore, the new, RMCM follow up queries might be:

For (B): Which of following animals is a feline animal?

For (C): Which of following animals is a small animal?

For (D): Which of following animals is a canine animal?

This approach indeed ratchets up complexity of the question. But, with reason: We can be confident that a student who can answer this RMCM question *knows* his animals and did not pick the answer to the original question at random.

This simple example illustrates when, where and how to use Reverse Multiple-choice. The original question of this example may have been aimed at elementary school students, and they may not have the vocabulary or maturity to come up with the reverse queries. But the same question and modified queries may be appropriate for older students, particularly for English Language Learners. RMCM items tend to be harder. And, not every multiple-choice question is useful to generate RMCM items.

5. Implementation Notes for RMCM

5.1 Question Creation: Expectations from an Examiner

While RMCM can reduce the burden on the examiner by mechanizing grading, much of the onus for effective use of the method is on the examiner.

A Reverse Multiple-choice question may be written *ab initio*, or created from an existing, imported multiple-choice question. In either case, an examiner has to put in forethought to examine suitability of the narrative stem as well as the answer choices of the question. Completely irrelevant answer choices may fit in traditional multiple-choice questions as "easy" but be totally inappropriate for query reversal. Examiners must also decide the purpose of each answer choice in terms of what is tested.

5.2 Context, Fact Objects and Fact Values

Implicit in *any* type of question and answer is the background or context of the narrative for underlying matter.

A rarely mentioned but notable merit of multiple-choice format is that it captures interpretative context *better* than other testing formats, for the simple reason that it is possible to view each putative answer of a multiple-choice question as adding contextual information for the interpretation of the narrative.

Furthermore, the selection of one answer choice as "correct" over the other choices generally turns on a few syntactic elements, such as, words, phrases, operations, numbers and symbols etc. In RMCM terminology, the syntactic elements that make an answer choice correct are called Fact Objects (FO). In our example of animals, the qualifier for the word "animal" is a fact object.

The value (perhaps a character string) of a fact object in an answer choice is Fact Value (FV) of the fact object, a concept akin to assignment of a constant value to a variable in algebra. In our example, farm, feline, small and canine may be regarded as the Fact Values of the qualifier fact object for the four answers.

When creating a question, the examiner specifies fact objects and fact values for *all* the answer choices. The system provides the platform and editorial support for question creation, later uses examiner's specifications to *automatically* evaluate student answers or flag unexpected student answers for human evaluation.

In a test taken by very large number of students, the proportion of answers flagged for human evaluation would be relatively small if the question is well constructed and appropriate to the students' academic level.

5.3 Question Creation Administration: Mode and Type of RMCM Question

Mode and type information for a question is specified by the examiner, assisted by the system in administrative functions and in generating reverse multiple-choice tasks.

Under RMCM regime, a question may be posed in three possible modes: A mode wherein the student is asked to select an incorrect answer, then asked to provide a modified or follow up query for which that answer would be correct; another mode wherein the student is asked additionally to identify the correct answer to the original query; and the third mode of traditional multiple-choice questions where only the correct answer is identified. This last mode is intended for administrative purposes of integration with standard multiple-choice tests.

The "type" of a question refers to the form in which the task of modification of query is specified for the examinee. There are many different *types* of RMCM question.

For each answer choice that the test-taker regards as incorrect answer to original query, such tasks include one or more of the following: identify the fact objects that need to be changed; identify the fact objects to change from a given list; write in the fact objects that need to be changed; write in the fact objects to change from a given list; identify the fact values of a fact object that need to be changed; identify the fact values to change from a given list; write in the fact values of a fact object which need to be changed; write in the fact values to change from a given list.

The write-in type questions are similar in spirit to "fill-in-the-blank," and identification type to "matching" questions. But the creation of such question in RMCM protocol is systematic and offers potential advantages of automation at various levels.

We note that graders of classical open format questions typically look for segments in the nature of "fact object" and "fact value" in a student's answer; Reverse Multiple-choice paradigm makes it possible to deploy computers.

5.4 Answering RMCM Question: Expectation from Examinees

In purely formal terms, for a test-taker answering RMCM questions may be a bit unfamiliar at first but not much harder than answering multiple-choice questions, and probably easier than answering long or short essay type questions.

The RMCM task may be specified in simple, familiar terms. For example:

"Find the words /phrases/ symbols or other segments of the query which, if they are changed, will make your selected incorrect answer the correct answer for the changed question."

Or,

"Your selected answer is incorrect because at least one query segment has the wrong value; identify which value from the given list should be assigned to the query segment(s) so that your selected incorrect answer becomes the correct answer for the changed question."

Similar language may be used for write-in answers for fact objects or fact values.

Answering Reverse Multiple-Choice questions, however, puts different kind of pressure on students used to traditional multiple-choice questions – learning the subject matter in-depth. Students must acquire the skill to deconstruct and reassemble a question, and learn to focus on closely reading the fact pattern, critically evaluating the answer choices and recognizing the critical pieces of information in the fact pattern on which the answers turn.

In other words, the students must be prepared for *in-depth* assessment. We may also note that RMCM format is *less prone to cheating* by test-takers than standard multiple-choice format.

5.5 Grading

The examiner's up-front work in question creation pays off in grading, by shifting a lot of the drudgery and cost to mechines or to assistants who cost less.

5.6 A Detailed Example

Q. Divide and simplify the expression: √(10) ÷ √ 2 =

(A) √ 8 (B) 2√2 (C) √5

(D) 2√5 (E) 2√3

The Correct answer is (C).

Here RMCM would task the student to modify the question suitably to make one or more of incorrect answers correct for the changed question. Thus, the incorrect answer:

(A) would be correct if we: change the division "÷" to a subtraction "-" and drop the second square root operation "√ **"**as well as extend the first "√ **"** to cover both the numbers;

(B) would be correct if we: change the "÷" to a "-" drop the second square root operation "√ **"**and extend the first "√ **"** to cover both the numbers;

(D) would be correct if we: change the division "÷" to " * " a multiplication; and

(E) would be correct if we: change the division "÷" to an addition " + " and drop the second square root operation "√ **"**as well as extend the first "√ **"** to cover both the numbers.

This question is designed to test as well the observation that the answer choices A and B are equivalent; hence the collated information for answer choices A and B above is identical.

The list of segments/syntactic elements here may be listed as: the first **√,** the **first paren (,** the **first number,** the **second paren),** the **arithmetic operation,** the **second √,** the **second number.**

A significant point that we note is that Fact Objects are *semantic* entities encapsulated in syntactic elements. Context for interpretation of the narrative of the question imbues it with meaning, but in written language *syntax* is the bridge over which meaning is communicated.

Answer choices of a multiple-choice question append the contextual lexicon along with the query narrative and call of the question in a natural, efficient manner. By emphasizing the role of Fact Objects and Fact Values, RMCM can utilize this strength of the format for in-depth testing.

In practice, this may be achieved by calls to databases and table look ups, etc.

Therefore, the system may maintain several related tables, for example, the illustrative Table 1 below for compilation of Fact Objects and Fact Values; as well as, Table 2 for grading purposes that shows fact objects/fact values for answer choices with the score breakdown. In a question where examinees are asked to identify or modify the fact objects for a given answer choice, fact object selection score column of Table 2 is relevant. When the question asks for the fact values for one or more fact objects, fact value selection scores are relevant.

Tables similar to Table 1 and Table 2, below are expected to be provided by examiner when making the questions.

TABLE 1. Fact Object/Fact Value Table

Fact Object	Ans. choice	Fact Value	Comments
[the first square root], √			No change for any answer
[the first half paren], (			No change for any answer
[the first number]			No change for any answer
[the second half paren],)	(A) (B) (E)	*(M, move symbol to location after second number)* *(M, move symbol to location after second number)* *(M, move symbol to location after second number)*	
[arithmetic operation]	(A) (B) (D) (E)	- [subtraction] - [subtraction] * [multiplication] + [addition]	
[the second square root **],√**	(A) (B) (E)	*(D, delete symbol)* *(D, delete symbol)* *(D, delete symbol)*	
[the second number]			No change for any answer

TABLE 2. Answer choice/Fact Object/Fact Value/Scoring

Ans. Choice	Fact Object	FO selection score	Fact Value	FV selection score	comments
(A)	**[the second half paren],)** **[arithmetic operation]** **[the second** square root **],√**	25% 50% 25%	*(M, move to location after second number)* **- [subtraction]** *(D, delete symbol)*	30% 40% 30%	
(B)	**[the second half paren],)** **[arithmetic operation]** **[the second** square root **],√**	25% 50% 25%	*(M, move to location after second number)* **- [subtraction]** *(D, delete symbol)*	30% 40% 30%	
(D)	**[arithmetic operation]**	100%	*** [multiplication]**	100%	
(E)	**[the second half paren],)** **[arithmetic operation]** **[the second** square root **],√**	25% 50% 25%	*(M, move to location after second number)* **+ [addition]** *(D, delete symbol)*	30% 40% 30%	

6. Classroom Experience

The following is a simple example from a quiz in an introductory Statistics course.

Q. (i) The common intelligence quotient (IQ) scale is Normally distributed with mean 100 and standard deviation 15.

What proportion of population has IQ scores between 115 and 130?

A. 68%. B. 95%
(C) 13.5% (D) 34% ?

(ii) Answer choice _____ is incorrect because _____.

(iii) Select one incorrect answer choice given above. Then change the question so that your selected answer choice is the correct answer to the changed question.

Herebelow is one student's answer to this question:

[Answer choice] B is incorrect

• B is incorrect because to cover 95% of the graph the IQ score has to be between 70 and 130.

[New question]

• The common intelligence quotient (IQ) scale is normally distributed with mean 100 and standard deviation 15. What proportion of population has IQ scores between 70 and 130?

Such questions throughout the course proved helpful to the examiner for grading since the job of matching answers against model answers was carried out by an assistant. Also, in cases where the students did not select the correct answers, the one-line reason and the changed question gave a window into the students' thinking and helped give partial credit where due.

7. Conclusion , Direction for Further Work and Research

Reverse Multiple-choice provides a viable assessment methodology, and an economical alternative to standard multiple-choice and open format testing techniques. It is particularly useful in situations where it is necessary to evaluate the extent to which the test-takers have mastered the art of recognizing critical pieces of information in the subject matter tested.

RMCM questions tend to be harder for the students, but their complexity deters both, mindless filling of the answer bubbles and cheating. These questions are harder for an examiner to create as well, but they can be reused over a longer shelf life.

Construction of the create and answer question platform is on its way to completion. We have planned an item library of RMCM questions for several subject area at various levels that could ease examiners' burden.

8. References

[1] I. M. Anand, "Method of developing educational materials based on multiple-choice questions," United States Patent No. 7033182, 2006.

[2] I. M. Anand, "Method of developing educational materials based on multiple-choice questions," United States Patent No. 8195085, 2012.

[3] Cathy N. Davidson, "Standardized tests for everyone? In the Internet age, that's the wrong answer," Washington Post Opinion, September 23, 2011

THE DEVELOPMENT OF TECHNOLOGY OF CREATION AND OPTIMUM PLANNING AND DISTRIBUTION OF RESOURCES OF INTEGRATED EDUCATIONAL PORTAL

The 2014 International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE'14), July 21-24, 2014, Las Vegas, Nevada, USA

Seilkhan Boranbayev
L.N. Gumilyov Eurasian National University
5 Munaitpasov Street
Astana, 010008, Kazakhstan
sboranba@yandex.kz

Bahtiyar Mukashev
L.N. Gumilyov Eurasian National University
5 Munaitpasov Street
Astana, 010008, Kazakhstan

Askar Boranbayev
Nazarbayev University
Astana, Kazakhstan
aboranbayev@nu.edu.kz

Sanzhar Altayev
L.N. Gumilyov Eurasian National University
5 Munaitpasov Street
Astana, 010008, Kazakhstan

***Abstract*—The article describes the developed technology and methodology of creating of an integrated information management system learning portal, its architecture, graphical design and functionality shell. The development of the informational portal's soft implementation is described also. Optimization of information portal architecture and the used technology to provide fast access to resources of the portal is described. The authors have covered multiple aspects of the design, development, implementation and maintenance of web-portal applications.**

Keywords—portal; information system; application; implementation; design

I. INTRODUCTION

Information portal - web-site which is giving direct access to resources, applications and services for demand thematically saved resources. Typically, portals provide search services, directories, forums, feature articles and blocks provide personal access that opens additional opportunities for users. Often these portals are called web-portals. [1]

Regard to the education information portal is seen as a multifunctional tool of learning in high school, which has significant educational potential, the implementation of which ensures the formation of information competence of students in the information environment [2].

Development of an information portal for distance education includes the execution of the next set of tasks:

1) Design and development of internet resource allowing 24 hours a day, 7 days a week remote access to educational materials.

2) Selection of portal's hardware and software platform.

3) Development of graphic design, database design and functional shell of the portal.

4) Setting up and implementation distance learning system

Designing user interface and the administrative control panel of the portal.

II. TECHNOLOGIES USED

In the path of solving this problems integrated system management portal has been developed. Integrated Management System portal is based on the use of technologies, the main ones are [3-6]:

- Platform-independent interface CGI
- Language development of PHP scripts
- Technology of interactive documents DHTML building
- ActiveX controls
- Server technology Active Server Pages (ASP)

Optimal platform for the implementation of technological and operational framework of the project is a bunch of Apache

+ MySQL + PHP under Unix-based OS. Control Panel's web-hosting is implemented using DirectAdmin.

III. WEB-SERVER' SOFTWARE

The server operating system is GNU / Linux (distributions CentOS, Red Hat, Fedora, Debian), which is a complete multi-user multi-tasking operating system. Most of the freely distributed on the Internet for UNIX programs can be compiled for LINUX almost without any changes. In addition, all the source code for Linux, including the kernel , device drivers , libraries , custom software tools are distributed free . Linux provides a complete set of TCP / IP protocol for network operation. It include device drivers for many popular card Ethernet, SLIP (Serial Line Internet Protocol), PLIP (Parallel Line Internet Protocol), PPP (Point-to-Point Protocol), NFS (Network File System). The complete range of services to clients and TCP / IP, such as FTP, telnet, NNTP and SMTP. Also , Linux is working with the software : MySQL 5.0.51a, Dovecot 2.0.13, Exim 4.76, PHP 5.2.17, Perl, BIND, Apache 2.2.18.

IV. THE STRUCTURE OF AN INFORMATION PORTAL AND ITS FUNCTIONALITY

The structure of the information portal includes three elements that form an integrated system of distance learning. Components of the portal: archive file's (it contains forum "question - answer"), a news blog, and immediately e-learning system.

Each of the three components is selected domain, so to increase speed of operation and reduce the time portal response from the database were created 3 database, which will work with either singly or, if necessary, simultaneously.

The general diagram of the database information portal for distance education consists of three databases:

1. Database news blog enu-info-blog consists of 16 tables.

2. The database file archive enu-info consists of 130 tables

3. Database of distance learning education consists of 250 tables.

Statistical result of loading the portal as a result of the separation of the portal database (the average rate of the portal enu-info.kz):

- Downloads page: 0.24 KB / sec (1.9 kbit / s)

- Access to the site: 0.58 KB / s (4.62 kb / s)

- Test of downloads: 3

Visualization is the final stage in the complex process of designing and beyond that is usually the most vulnerable point in creating an information system. Since the interaction with the target audience, or goes directly to the first through the GUI.

Informational portal for distance education is structured into three parts, each of which performs a specific task. Due to the specifics of each of the portal element for each of them have been developed various types of graphic design.

Logo - one of the main visual attributes of the portal, his face, so it should reflect the direction of its activities, to be simple, memorable and competitive ability. For the logo must meet certain requirements. Adhering to these requirements, combined logo for portal was designed..

Fig. 1. The Logo of the Portal

User interface information portal for distance education carried out in accordance with the standards ISO 9241-12-1998 « A visual representation of information», ISO 9241-14-1997 «Menu», ISO 9241-16-1998 «Direct manipulation », ISO / IEC 10741 - 1995 «Cursor », ISO / IEC 11681-2000 «Icons» and complies with all regulations and guidelines in the field of web- design.

The portal is implemented dialogue with the user before any manipulation associated with a change or receive information. Passwords and personal data are stored in separate tables, which only has access to the system and, in extreme cases, the administrator.

User becomes a full member of the portal for distance education and after going through the registration confirmation.

In the event of a problem with your user account password recovery service is provided login and password to the specified email address.

After logging in the portal user becomes a full member of the authorized distance learning. It can edit their personal information, to download their data or other users of the portal to add them to friends.

Distance learning system information portal also includes a user interface over an extended character. Here the user has the rights package for the manipulation of his public profile, in accordance with its role in the portal (administrator, student, course author, teacher, etc.).

Profile information may be available in extended and shortened version, there are mandatory and optional fields in the profile.

The portal operates online service correspondence between authorized users, messages are stored in a separate archive for each member of correspondence.

Also available file storage service on the portal server for this is implemented special user interface to add and delete files.

A separate item profile of distance learning is the development of a module multilingual information portal. Each user has the ability to choose from a list of suggested language preferred language for the menu and directly for the most training. For the convenience of lectures and courses, the home page of the portal is fully customizable, menu items, profiles and find how you can hide the tab bar.

Information architecture - a concept that is closely related to the content, structure and optimization. Under the system architecture refers to the organization of files, pages, and other information contained in the portal. The main purpose of information architecture of the website - is to create a resource that would allow the user to quickly find the information you need.

As already mentioned, the structural portal consists of three elements: a news blog , file archive and distance learning system . To increase the maximum comfort and ergonomic web- portal architecture of the portal was also divided into 3 elements, which together form a single system. The system successfully solves the problems outlined concepts due to the CDN (Content Delivery Network) - a geographically distributed network infrastructure that allows to optimize the delivery and distribution of content to end users on the Internet. Using a CDN increases the speed of transmission and users to download audio, video , software , gaming and other digital content.

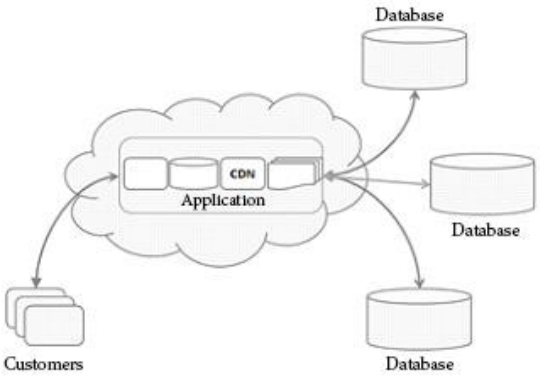


Fig. 2. The structure of the information architecture using

The optimal scheduling and allocation of resources portal includes optimization of the information architecture of the portal, organize pages, the distribution of internal links, affixing different tags needed for search engine promotion. This is achieved using this process-oriented content management system, which provides a process of systematization of all the content of the portal.

Each article belongs to a particular section, each section is structured in a certain category , the optimization process is achieved by assigning each structural unit of the hierarchy of the individual number which is entered into the registry database. Related articles from a common kinship categories are entered in a table , and to optimize the search, and download time required web- content .

Implementation of advanced intelligent search content on the portal is made separately for each component of the project. So, your search is as file storage , a forum, a news blog , and distance learning system . The format of the search is performed in 2 ways, either by direct input on the web- search page, or through the form on the home page.

Content Management System (CMS - Content management system) - is an information system that is used for the organization and collaborative process of creating, editing, and content management portal . The main purpose of this system is the ability to gather together and unite on the basis of the roles and tasks of all heterogeneous sources of information and knowledge available within the portal, and beyond, and the possibility of interaction between the portal administration and working groups established by the project with their bases knowledge, information and data so that they are easy to find , remove and reuse familiar to the user.

Portal is structured in three elements were created basic system that supervises, and systems for the control element by element. The basic system is installed on the main domain information portal system control element by element - the subdomains of the portal.

So, we have the technological and functional basis of the information portal configured using 3 content management systems: Joomla, Wordpress, Moodle.

All three systems are selected on the basis of compliance with the majority of the selection criteria CMS. Graphic templates and databases have format and encoding, so the layout and the transfer does not have any problems. CMS system meet the minimum system requirements of the server: PHP 4.3.10 or later; MySQL 3.23.x or later; Apache 1.3.x or later.

The whole concept of software development portal includes work on the creation: modules, components, plug-ins and widgets.

The basic software modules of the portal:

1) Navigation Modules - main menu and chapter menu , which can consist of graphic buttons or text links . Perhaps the use of pop-up when you hover the mouse cursor over the menu sections. Uses redundant modules navigation: back button, links to the next in the structure of sections and subsections, sliders, site map, drop-down lists of rapid transition. Duplicate navigation can significantly reduce the search time and reference time of the decision.

2) The form of authorization and restriction of access is used as a tool to control visitor access to a variety of materials from the portal, shares registered users, customers, partners and guests.

3) The posting form will allow users to send all sorts of messages directly from the portal to the e- mail address or to database without resorting to the use of e-mail programs. These

forms can offer the visitor to choose the destination mailbox. The posting system equipped with anti-spam.

4) Form newsletter subscription and automatic distribution module allows the guests and partners to sign up to receive news and special offers portal and the portal can be sent to many addresses news, offers , promotions and information on new publications .

5) The search form on the site allows visitors to easily find information of interest on the portal. Formation of the search reports are already going to the laws of search. Used morphological analysis of search queries and the ranking of links for relevant queries

6) Polls, surveys, questionnaires allows visitors to participate actively in the work of the portal, in search of strategies to attract new users.

7) Guestbook, Forum - This is the section where every guest will be able to express their wishes, advice, criticism, review , opinion, ask their questions .

Each module is mounted on the portal has a clear position and dimensions within which it is located, the right to control the parameters of the modules has the administrator through the administrative panel of content management system.

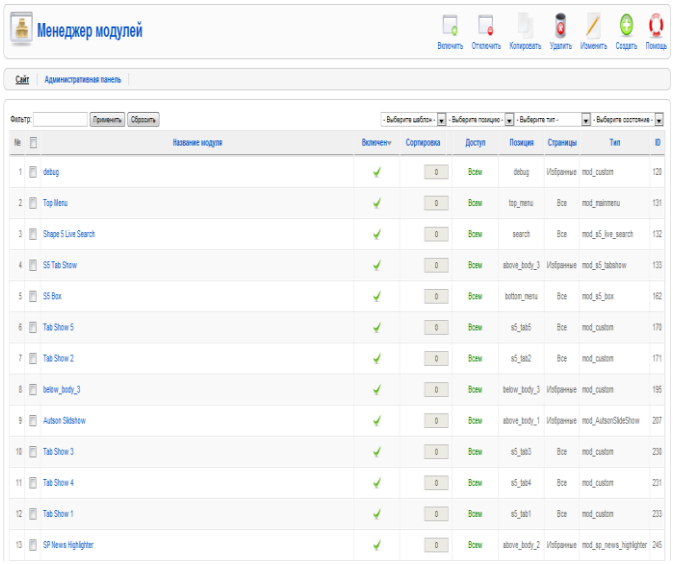


Fig. 3. Modules Manager Panel

On the server, each module is stored in a folder with the executable file after the transfer module in the CMS assigned a number and the level of access to the module (or is accessible to all , or only for registered users). The main module list includes 60 names. Each module is implemented only for the finalization of the functional capabilities of the portal.

Component - this is the main means of enhanced functionality and is a set of scripts. Components are the forums, file archives, galleries, statistics collection system, backup, etc. They also perform: user management, partitioning, addition and display materials, etc.

System supports basic and auxiliary components. The main components implement forums, hit counters, file server portal, web- player output. Auxiliary components installed in the CMS, output voting, news, search, directory of links to the main panel of the portal, etc.

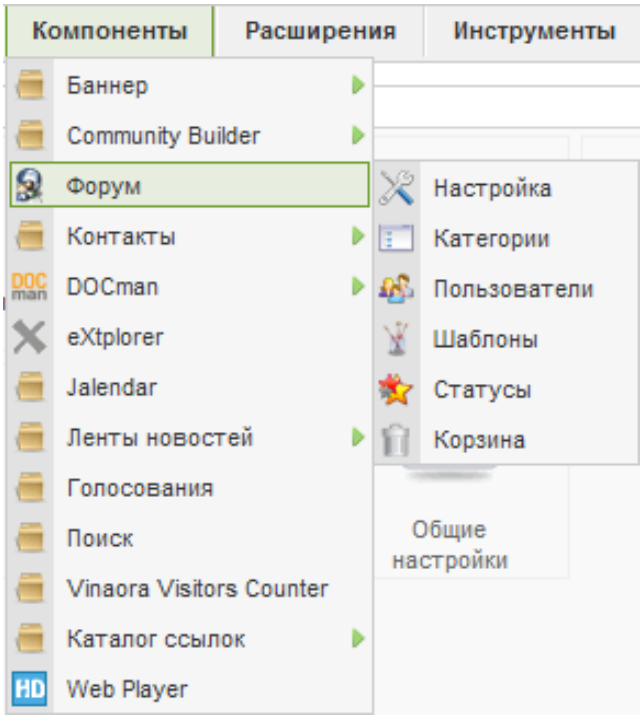


Fig. 4. Installed components Menu

Also, each component has a set table in the database of CMS system and limit access

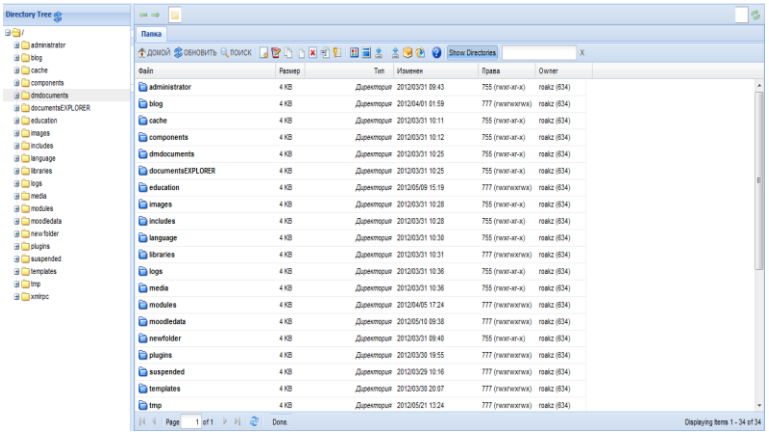


Fig. 5. Portal's file archive component

Plug-in - a function that performs some manipulation of the data before displaying them to users. The group includes a plug-in editors visual editors. Visual editors greatly simplify the process of adding materials. The portal default plugins operate some modules and components that allow you to implement these extensions directly to the portal.

The main attention was paid to the optimization of social life , so in order to promote the portal in social networks have been

implemented social tabs on the article and go to the window to the community in various social networks with plugins.
Another building block of the functional basis of the portal is a widget, it is installed and is designed specifically for the CMS Wordpress. Widgets are used to add a variety of portal functionality. The management process is carried out by means of widgets admin panel.

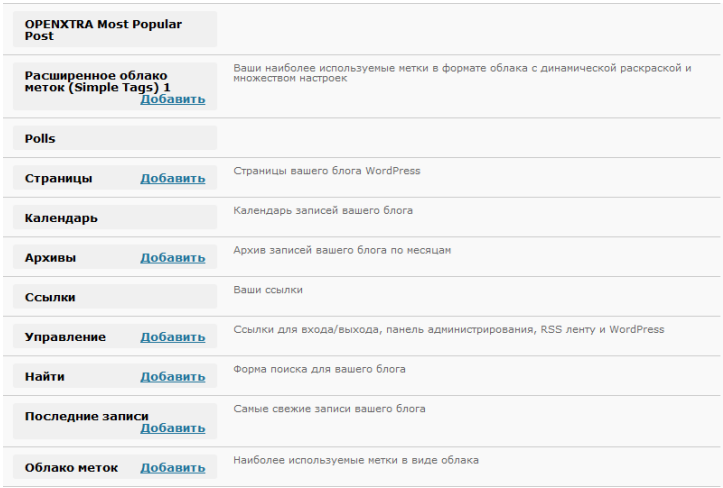


Fig. 6. News blog widget's Menu

Finally, in order to expand the functional basis portals were developed 70 modules, 12 plug-ins, widgets, 11, also used on the portal modules, plugins and components installed in the default CMS. The general scheme of the functional information portal combines the structure consists of four elements: modules, components, plug-ins and widgets. The successful operation of each element to successfully operate the system.

REFERENCES

[1] Budaev E.S. Information technology training and development of analytical problems on the web-portal. – Moscow. - 2007. - 228 p.

[2] Tikhonov A.N. Internet portals: content and technology. - Moscow: Education. -2007. - 720 p.

[3] The terms of web-technologies, platform-independent interface CGI: http://webstroymast.narod.ru/words.html

[4] The programming language PHP: http://php.find-info.ru/

[5] Dynamic HTML: http://www.intuit.ru/department/se/pinform/6/

[6] Using the Controls ActiveX: http://wm-help.net/books-online/book/56472/56472-90.html

E-Learning Through Social Networking Sites-Case Study Facebook

K. Hadjar[1] **and K.R. Hammoud**[1]
[1]Department of Multimedia Science, Ahlia University, Al Manama, Bahrain

Abstract - *Online learning has been acknowledged as a successful and useful pedagogical technique and tool. Thus, it is widely incorporated into a wide variety of teaching and learning methods in higher education. The use of Virtual Learning Environment (VLE) in higher education has become an essential method for good education. This field, however, has not been the topic of thorough research when it comes to online learning, delivery and evaluation. This research tackles education from an online learning point of view, mainly through the focus on social networking channels and integrated applications. It puts forward an integrated outline over the most employed and spread social network – Facebook – with three key components for online learning via VLE.*
Keywords: E-Learning, LMS, VLE, Social Networks, Facebook.

1. Introduction

Facebook [11] and other social networking tools, have increasingly become the object of scholarly research. Scholars in many fields have begun to examine the effect of social-networking sites, exploring how such sites may affect the issues of identity, privacy, social capital, youth culture, and education. They have been greeted and encouraged for they are capable of offering new, socially engaged educational experiences for students and learners in undergraduate, self-directed, and other educational sectors.

A theoretical and historical analysis of these Social Networking Sites in general, and Facebook in particular helps to locate and qualify this perspective [1]. The analysis of dominant media in particular presents questions as to whether social networking sites' platforms can satisfy the development of the capacity for debate and disagreement which is a critical factor of learning. Several websites are beginning to go deeper into the power of the social networking model. These models provide a means for connecting fragmented industries and small organizations lacking the resources to reach a broader audience with interested users. Social networks are striving to offer a special way for individuals to communicate digitally. The sharing of information and ideas is made possible through these communities of hypertext. This old concept is now placed in a digital environment.

Facebook has the ability to offer a great platform for education. Students use Facebook on a daily basis for many interests of many kinds. One of the best ways to use Facebook as a Learning Management System (LMS) is through groups. Many previous researches show that this idea many benefits for learning and education in general [4]. Furthermore, using learning management systems can encourage student enrollment and can also create interaction between learners and instructors as well. In the case of Facebook being used as a Learning Management System, it would help solve some problems faced in education. Furthermore, a Facebook application, regardless of its simplicity or globalism, can replace Blackboard or Moodle or any other LMS with its simplicity and accessibility. One can easily access a Facebook application from a variety of electronic devices like computers, PDAs, Tabs, and even Smart Phones.

Creating an e-leaning system through Facebook' social networking site is the object of this paper. The key objective of this research is to introduce the e-learning through Social Networking Sites as an innovative and modern way for communication between academics and scholars [9]. Other objectives lie in proving that traditional e-learning present systems, taken up by many education institutions and universities are not sufficient nowadays in a world shifting towards being open in every possible way. The inadequacy of the traditional e-learning software (Moodle, Blackboard) [5] lies in their failing to be local applications. In a world where learning is becoming almost free of all charges, and depending on a low cost internet connection, the ability of being online is an obligation for any new e-learning application to be introduced. Besides, the ability for learners and researchers to publish their works in the easiest available way is of equal importance. This paper is organized as follows: in section 2 we discuss the impacts of adopting the ICT within Education. Section 3 presents Social Networking as a concept and how it can be integrated within e-learning. In section 4, our model named EduSocialNet is exposed in detail. Finally section 5 concludes this paper.

2. Education and ICT

Many researchers have shown that online delivery can produce creativity, high order thinking, reflection in action, and skill proficiency. It can even boost these educational values through active and engaging learning experiences. Furthermore, it has been argued that ICT assist problem- and task based learning so that they get more communicative and interactive; this way the restrictions of face-to-face learning can be overcome. This produces various learning styles such as blended, interactive and innovative learning models. In addition to this, it has been argued that this form of learning gives students the opportunity to learn according to their preferred learning techniques. This way they can become more self-directed and more responsible for their learning. In other words, ICT approaches give way for new forms of interactions that are flexible, personalizable and customizable.

In spite of the fact that ICT has progressed sufficiently to allow educators to exemplify many features of face-to-face learning in an online learning environment, researchers insist that it still has restrictions when it comes to being functionally capable of replacing face-to-face learning. For instance, the design teachers' demonstrations of a development procedure of a given product or

technique may be limited to real-time delivery and interaction due to technical limitations and complications such as bandwidth limitations and content formats for internet browsers. ICT-based learning environments need to be customized to suit the various learning styles and modes, particularly focusing on the communication efficiency as well as learners 'participation and commitment. In addition to this, it needs to improve technical limitations in order to reproduce characteristics of a discipline and achieve pertinent educational values.

2.1 Virtual Learning Environments

A Virtual ―Learning Environment (VLE) can be defined as a flexible e-learning and online community system for delivering online courses and establishing online communities. A VLE has a dual role; it can be used either to improve traditional class formats or deliver a whole course online. For instance, asynchronous discussion boards are used as communication channels where students can take part any time and from any place where they have access to internet or intranet connectivity. More generic features of VLE are the following: the instructor can utilize discussion boards; continue class discussions outside the classroom‖; ―encourage an online community; expand group or individual student facilitated discussions; post and discuss case studies; post student papers for peer evaluation and critique; post homework questions; offer a public opportunity for students to post questions‖; ―provide a medium for a guest speaker Q & A; and create an online social forum for a course (blackboard [12], moodle [13]). There exist also additional applications available such as grading, surveys, statistics and announcements that are useful in terms of improving and supporting interactive learning in VLE. The ability to arrange various applications and tools in VLE is crucial for instructors to shape the unit delivery and format according to the unit objectives.

3. Social Networking

Social networking is known as an online service or platform that has as its main focus the creation of social networks based on social relations between people who share the same interests or activities [2]. In some other cases, those people may be sharing the same problems or backgrounds, or also sharing a real life connection. The social networking service is characterized by a profile created by a user, sharing personal or general information about himself, with the presence of links to other additional services [11]. The majority of social networking sites or services is web based and offers a variety of services for their participants in terms of communication services like e-mail, instant messaging, posts, chat, tweets …

The most known types of social networks are the ones containing categories like a school year or classmates of a specific school or university. Others provide ways for friends to connect and share photos, videos, ideas, and memories … [6] The biggest and most known social networks or our era are: Facebook, Twitter, Google+ and Hi5 … they are being used all over the world and by people from all age groups [7]. Although online social networking has many risks, it also has many potential benefits. Social networking can offer opportunities for new relationships as well as strengthening existing relationships, whether your kids 'friends are close to their homes or across the world [3]. Even though it is good to encourage positive relationships through various avenues, including the Internet, it's important to be cautious when your kids are getting involved in online social networking.

Social networking is becoming increasingly important in schools. Facebook, Moodle, Blackboard and other sites are frequently used by teachers to communicate with learners or for out-of-classroom discussions. Youth can further investigate topics that are of interest to them through online social networking. By making connections with other people who have the same interest, kids can learn and exchange knowledge with those they might not have had the occasion‖ to interact with [8]. Teachers often take advantage of students 'social networking abilities to create class blogs, discussion forums, videos, and more. By collaborating with other students and teachers through online social networking, children are capable of‖ building stronger school communities. A social networking site can be a fine way to make connections with people with similar interests and goals. They can be a way to connect with or "meet" people that a student may not have had the opportunity to before including other students, staff, faculty and even alumni.

The benefit of use of Social Networking for learners can be:

- Young people as social participants and active citizens: Social networking services can be used to manage activities, events, or groups to showcase issues and opinions and make a larger audience‖ aware of them.
- Social networking services rely on active participation: users participate in activities and discussions on a site, and upload, alter or create content. This helps supporting creativity and it can also support discussion about ownership of content and‖ data management.
- Social networking services rely on active participation: users participate in activities and discussions on a site, and upload, alter or create content. This helps supporting creativity and it can also support discussion about ownership of content and‖ data management.

Social Networks present some disadvantages dealing with security despite the efforts being made by Social Networks companies to improve security. In fact there are some cases of harassment, cyber-stalking, online scams, and identity theft noticed among Social Network's Users [10]. Also the posting of information among the social networks 'sites needs to be verified by a moderator in the context of e-Learning to avoid disseminating wrong information.

Despite the disadvantages we still believe that e-Learning through social networks helps overcome many shortages that we find inside LMS.

3.1 Study and Findings

In order to see if a Social Networks can be adopted as LMS, we did a study. We created a Facebook group as an LMS and monitoring students 'observations, in one of our university courses, and comments based on their experiences. The results and findings confirmed that this kind of groups has enough potential and can be used as an LMS. It enables students to share resources, make announcements, participate to online debates and discussions, and answer questions which are the basic functions of an LMS [8]. Using the Facebook group as an LMS allows instructors to have more control on their material and courses than other commercial LMSs, it also solves the problem of complexity of other commercial LMSs as well. However, this study showed that using the Facebook as an LMS has a number of constraints as well.

The Facebook group does not support learning resources in other formats such as PPT or PDF to be uploaded directly, and thus third party web sites have to be used to overcome this limitation. In these two courses, Google Docs was used to host learning materials of other formats, and Kwiksurvey was used to collect feedback from the participants. The result showed that the integration of external tools supplemented and enhanced the capability of the Facebook group as an LMS. The Facebook group is not perceived as a safe environment, even though it provides different access control such as being open, closed or secret. In this study, the Facebook group was set to closed, and the students were not required to be friends.

The positive result of this study implies that the Facebook group can be used as an LMS substitute or supplement. In schools where commercial LMSs cannot be afforded, the Facebook group can be used as a fully functioning LMS. In other schools where commercial LMSs are already in use, the Facebook group can be used in extracurricular activities to supplement social interactions and personal profile spaces, which are often found insufficient in LMSs.

This study faced some limitations. The students felt insecure in this study as they used existing Facebook accounts. Research shows that learners tend to separate life from studying and home from lectures. They do not want to mix learning with social lives. Future research will explore student perceptions on using a separate Facebook account for learning. However, it may lose social dynamism and also may compromise the kind of sustainability that makes Facebook popular.

This study did not compare the effects of using Facebook as an LMS with other commercial systems like Blackboard. The causes and effects of using Facebook to support learners 'learning in blended or online courses were also not investigated.

The Facebook group has the potential to be used as an LMS. It has pedagogical, social and technological affordances, which allow putting up announcements, sharing ideas and resources, and implementing online discussions.

Using the Facebook group as an LMS, however, has some limitations. It supports almost all formats files to be uploaded directly into the group, but it doesn't support reading and editing them all, it is limited to text, tables, and presentations in general. A Facebook group has to be monitored at all times if adopted as an LMS in order to provide a safe and appropriate learning environment.

4. EduSocialNet: our proposed model

A model is being created showing the necessary stages for creating an online system where academics, including instructors, students, and researchers, can communicate in a public controlled space with all the needed facilities to maintain their work and share it with the appropriate groups and communities. Our Model will have a complete scenario that will describe all the operations that can be possible elaborated into the system.

The main concept behind this research is to adopt a virtual learning environment (VLE), based on a social networking site: Facebook One of the most significant advantages is that the wall app and Facebook in general, is accessible by every communication device that can be connected to the internet, which means that our integrated app will always be available for students participating in a certain course to post, edit, publish, discuss, upload, download, and share all kinds of ideas and materials concerning this course. All the sharing of all kinds of materials would be limited to the members of the same course only. Others should not be able to see, read, edit, or modify any of the contents posted by members.

On the other hand, the instructor for the course, assigned as an administrator of the group, should be granted the authority and the ability to add and remove members to the group. The instructor should also be able to do all the functions available for members; in addition, the instructor, as an administrator, should be able to delete any inappropriate content published on the course page.

The student's and instructors' functions of EduSocialNet are as follows.

Student's functions of EduSocialNet:

- Login, Modify own profile, Change own password, Join a course, Create post, Edit own post, Comment on others' posts, Modify own comments, Delete own comments, Publish a note, Publish a photo/image, Upload a file, Download a file, Send private message, Chat with other members privately

Instructor's functions of EduSocialNet:

- Login, Modify own profile, Change own password, Create a course, Join a course, Administer a course, Edit course profile, Invite/add students, Publish discussion subject, Publish course material download links, Publish own images, Delete own and others' images, Upload files, Download files, View/Edit course files, Access app database, Add own note, Edit own note, Delete own and others' notes, Create post, Edit own post, Delete own post, Comment on others' posts, Modify own comments, Delete own comments, Chat with other members privately

Regarding the Facebook group, public group are not good because everyone has access to it, whereas closed groups are more private. Only group members can see the members, posts, and all the posted materials in the group. On the other hand, the name of the group is still public, and can been seen by all the users of Facebook. Private groups are the ideal ones to use as a VLE. This kind of groups is secret, and cannot even be seen in the list of groups in Facebook.

Figure 1, contains a snapshot of the group creation.

Figure. 1. Snapshot of the group creation

At any given time or moment, the Administrator can deny anyone from posting in the group and keep this privilege to himself only.

Writing a post is the easiest task to do in a group, a student has to enter the text and press the post button and the text will be displayed to everyone in the group, depending on the permission granted to the members, some of them are not permitted to post, some others are not permitted to see others' posts. Figure 2, illustrates writing a new post on the group wall.

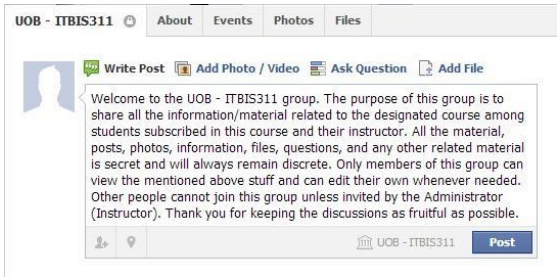


Figure. 2. Writing a new post on the group wall

Depending on the permission granted to a user in a given group, a student or Administrator can add a file to the group. Three sources for the file are available. The first source is to browse his computer and add the desired file. The second source is to connect his Dropbox to his Facebook account and gain the ability to choose his file from his online files. And the third option is to create the file locally on his Facebook screen. Figure 3, shows uploading a file to the group.

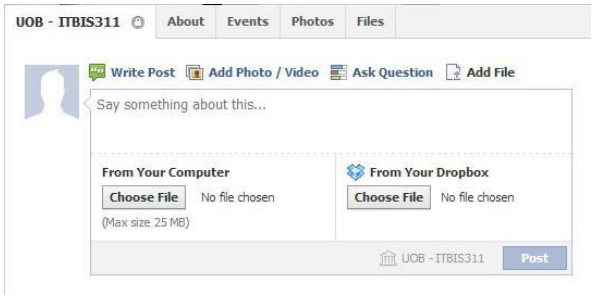


Figure. 3. Uploading a file to the group

One of the best features of the Facebook group is the ability to create events and share them. The instructor can create a new event, give it a significant name, insert all the needed details, specify the location, choose the start and end date and time, and configure the privacy settings for this event. See figure 4.

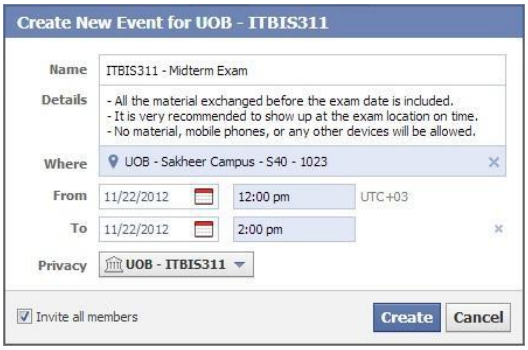


Figure. 4. Creating a new event for the group

The Facebook group used as an LMS will not be feasible without any backup solutions. Facebook have recently introduced the option of downloading all your data and save it as a package on any computer or drive of your preference. Backups can be taken per semester and archived on a local server on campus.

5. Conclusion

This paper presents our contribution to create EduSocialNet, our proposed model that adopts social networking as a LMS. Through our small testing we argue that Social Networking such as Facebook can be used as an LMS. It is actually being used nowadays by many instructors in the academic field in general. Universities have already started to show themselves on Facebook for numerous reasons including advertising and commercial reasons. Instructors, who have already begun to adopt this system, including ourselves, are doing it unofficially. That's mainly because many obstacles stand in the way. The most important one has to do with the legal issues that the universities may face in general. As for other issues such as privacy and security, they can be solved with an appropriate code of conduct and good practice. Our future work is divided into two main parts. The first part comprises a full design of a Facebook application that will have full access and control on the Facebook groups created by instructors in a specific university. It also comprises the creation of a local database on a local server, inside the university campus, that will host all the information related to all students, courses, instructors… The second part includes the creation of a link between the new created database and the existing registration system inside every university. This kind of links can be achieved through many kinds of software available in the market, like Microsoft Access, Microsoft SQL, or Oracle for best practice. A user friendly interface will complete the whole job the best way possible.

6. REFERENCES

[1] Greenhow C., and Robelia B., *Old Communication, New Literacies: Social Network Sites as Social Learning Resources*. Journal of Computer-Mediated Communication, 14, pp 1130-1161, (2009).

[2] Boyd M. D. and Ellison N. B. *Social Network Sites: Definition, History and Scholarship*. Journal of Computer-Mediated Communication, 13, pp 210-230, (2008).

[3] Williams D. L., Crittenden V. L., Keo T. and McCart P., *The use of Social Media an exploratory study of usage among digital natives.* Journal of Public Affairs, 12(2), pp 127-136, (2012).

[4] Velestsianos G., *Higher Education scholars; participation and practices on Twitter*. Journal of Computer Assisted Learning. 10.1111/j.1365-2729.2011.00449.x, (2011)

[5] Chao I. T., *Moving to Moodle Reflections two years later*. Educause Quarterly 3, (2008)

[6] Friesen N. and Lowe S., *The questionable promise of social media for education: connective learning and the commercial imperative.* Journal of Computer Assisted Learning, 10.1111/j.1365-2729.2011.00426.x, pp 183-194 (2010).

[7] Junco R., Heibergert G. and Loken E. *The effect of Twitter on College Student engagement and grades.* Journal of Computer Assisted Learning, 10.1111/j.1365-2729.2010.00387.x, pp 119-132, (2011)

[8] Dawson, S., *Seeing the learning community: An exploration of the development of a resource for monitoring online student networking*. British Journal Of Education Technology, 41(5) pp 736-752, (2010)

[9] Barnes, S. B., *Relationship networking: Society and Education*. Journal of Computer-Mediated Communication, 10.1111/j.1083-6101.2009.01464.x, pp 735-742, (2011)

[10] Waters, S. and Ackerman, J., Exploring Privacy Management on Facebook: Motivations and Perceived Consequences of Voluntary Disclosure. Journal of Computer-Mediated Communication, 17, pp 101-115, (2011)

[11] Facebook , retrieved from http://www.facebook.com

[12] Moodle, retrieved from https://moodle.org/

[13] Blackboard, retrieved from http://uki.blackboard.com/sites/international/globalmaster/

Moving Legacy Device Forensics to an Online Format

Professor and Cybercrime Training Lab Director Eamon P. Doherty CCE, CPP, SSCP
School of Administrative Science, Fairleigh Dickinson University, Teaneck, New Jersey
Ester G. Doherty, M.A.S, Fairleigh Dickinson University, Teaneck, New Jersey
Elly Goei M.A.S, Fairleigh Dickinson University, Teaneck, New Jersey
Doherty@fdu.edu, FDU, H-DH2-16, 1000 River Road, Teaneck, NJ, 07666

Conference : EEE2014

Abstract **-** In 2013, there are still many occassions when outdated media must be examined for insurance purposes or in corporate policy investigations or criminal investigations. This paper discusses the need for teaching legacy media forensics and gives some specific examples of media that needed to be examined. The paper also proposes a model of how the subject may be taught in person as well as online. The paper shows the reader a combination of technical topics, legal topics, and educational topics.

Keywords: Legacy Device Forensics, e-Forensics, Online Learning

1 Introduction

Some people say that the pager has been replaced by the smartphone and that pagers have become extinct. However; this is not true. Glenn Bischoff, a writer for "Urgent Communications," states that 3.5 million pagers are in use in the United States in 2012 in the public safety and health care industries [1]. Pagers often get a message almost instantaneously while text messages to cellular network customers take two to three minutes and are limited to 130 characters. A three minute delay in a message can make the difference or not in being able to save lives and property from the destructive forces of fire. Insurance investigators may sometimes ask to examine pagers to see what was known at what time and to learn when people responded. The restaurant industry also relies heavily on pagers. There is still a need for examining pagers.

An adjunct professor, who also a full time police detective, once told Dr. Dohery and a class at Fairleigh Dickinson University (FDU) that he arrested someone with a large piece of plastic containing pockets which held CDs with child pornography on them. The plastic was rolled out and hidden between a box spring and mattress. The story demonstrates the need to know how to examine a CD. If the CD is scratched and not readable, then it is important to know what tools are available to recover the text, videos, and graphics on a CD or DVD. Many people are also unaware that there is so much diversity in the formats of the CD and DVD. Here is a partial list of the formats that an examiner may encounter: "CD-i, VCD, SVCD, SACD, CD-ROM, CD-ROM XA, CD-R, CD-RW, CD-MRW, DVD-ROM, DVCD, DVD-RAM, DVD-R, DVD-RW, DVD+R, DVD+RW, DVD+MRW, DVD+R Dual Layer, DVD-R Dual Layer, DVD+RW Dual Layer, DVD+VR, DVD+VRW, DVD-VR, DVD-VRW, DVD-VM, DVD-VFR, BD-ROM, BD-R, BD-R DL, BD-RE, BD-RE DL, BD-R SRM, BD-R RRM, BD-R SRM+POW, BD-R SRM-POW, BDAV, BDMV HD DVD-ROM, HD DVD-R, HD DVD-R DL, HD DVD-RW, HD DVD-RW DL, HD DVD-RAM, HD DVD-Video [9] "

There are also occasions when old computers and media are put out for garbage and then scavenged by people looking for a free system for family members and friends. There was one story of this relayed to the professor in class. The student in a continuing education class was a policeman and said that he received a report from a scavenger who picked up a system next to the garbage with child pornography on it. The system was antique and had a 5.25 inch floppy drive and a large full height hard drive. A report was made and a search warrant was obtained for the address of where the system was obtained. When the search warrant was executed, a computer seizure and examination was performed. More child pornography was uncovered on a new system and some obsolete media was also seized. It becomes obvious that there is a need to be able to examine obsolete media as well as new hard drives and media. Let's suppose that some media was not able to be examined. Wouldn't it be possible that a defense lawyer may try to create a doubt in a jury by stating that there may be exculpatory evidence exist in the unexamined media? If one member of the jury has a doubt, then there may not be a conviction. This scenario reinforces the case for being able to examine all the seized digital media from the crime scene.

In 2009, Dr. Doherty was approached by a leader of a credit union asking for help in imaging and accessing the data on an eight inch floppy diskette. The disk was from the 1970s and was believed to be the sole sourceof information about the beneficiary to an insurance policy. After a great amount of unsuccessful effort, Dr. Doherty could not obtain the

hardware and software to perform such an operation and referred the man to Kroll OnTrack Systems [2]. It became obvious that the need for such training on how to recover that data was evident. In the fall of 2013, a continuing education student had said in class that he and his son were going to specialize in investigating legacy media and devices. The man had spent all his adult life in computing and his son is a defense layer who is looking for a specialty to increase his workload. Legacy devices and media were perceived by them to be a less profitable specialty that was overlooked by the larger investigation and law firms. The father of the defense lawyer said that the classes on various legacy media topics should be taught both in person and online. He said that in person classes were good to get the hands on practice. He also said that online classes that included short videos with supporting text would be good for his son who lives far away and often has to go to court on short notice.

In 2011 and 2012, Dr. Doherty and Elly Goei researched how to obtain data from 100 MB zip disks, 250 MB zip disks, and from jazz disks. They obtained the hardware for those disks and performed an examination of the data for a colleague. The data was then transferred to a USB flash drive and a verbal report of the contents was given. All these situations are just some of the examples of the need for teaching people about the policies, procedures, and special techniques for examining legacy devices and media.

2 Teaching Legacy Device Forensics in Person

Dr. Doherty created and taught a legacy device and legacy media class for the FDU Digital Forensics Special Interest Group which is made up of a wide range of people in the FDU community. The class was expanded from one hour to three hours and taught for FDU's Continuing Education Department. There were many people that had an interest but could not attend because of work commitments, heavy traffic, and the cost of travel. The class included a variety of hardware that was obtained from eBay or discarded by acquaintances of Dr. Doherty. The class was divided into sections. Each section discussed a type of media or legacy device, its characteristics, its history of usage, and a methodology of how to recover the data from it. There was also a discussion of the chain of custody, The Frye Test, Voir Dire, The Fourth Amendment, Search Warrants, and some legal concepts such as "Fruit of the Poisonous Tree."

3 Porting Specialized Classes to Mooc

An administrator at the school said that the class could easily be ported from an in person class to one that is online. The administrator said that the class would be a good candidate for being a MOOC. An online class or MOOC could follow the same format. The history of each legacy device such as the pager could be discussed, it peak years of usage, and the number of people still using such devices. Then the various models of such devices and the companies who made such devices could be discussed. There could be a couple pages discussing the file systems and URLs of places where online user manuals could be obtained. There could also be video showing the student how to use tools such as the Paraben Project-A-Phone to photograph screens. Many of the videos needed already exist on Youtube or SecurityTube so further production of video is not always necessary. The online instructor would only need to provide a link to an existing video with relevant quality content. If the tools and techniques for digital evidence seizure and examination were law enforcement sensitive, then the student would be referred to places such as Teel Technologies who teach a JTAG and chipping forensics class. JTAGs have to do with an interface within electronic devices that are used for quality assurance and testing circuits. They can also be used for probing chips and obtaining data [3].

The digital media can also be taught using the same methodology. A picture of the digital media such as the eight inch floppy drive is shown. The history of the media and its years of manufacture and peak years of usage is discussed. The formats, file systems used, and other pertinent data is discussed. Then the chain of custody and the preservation of evidence using Faraday Bags and shielding from heat and magnetic waves are discussed. Lastly tools such as DriveSpy by Digital Intelligence and other tools such as Access Data's Forensic Toolkit are discussed. Classic subjects such as data hiding and steganography are discussed using text and video.

Lastly, the types of hardware that can be interfaced with modern computer systems are discussed. There are many vintage computer hobbyists who have recently created hardware and customized software to help people adapt mainframe eight inch drives to the computer. The Shugart Eight Inch Drive was originally made as a peripheral device for IBM Mainframe computers [4]. However; there is now a power supply board and interface kit that provides most of what someone needs to connect an eight inch drive to a personal computer www.dbit.com. There are many types of eight inch disk floppy drives available for sale on eBay. There is also a lot of documentation and boards for sale online that allow people to connect these floppy drives to modern computers. The Mitsubishi M2896 - 63 - 02M, Caldisk 142M, Shugart 801, Shugart SA850, Shugart 851, The Tandem Eight Inch Drive, The Memorex 550 Eight Inch Drive, and the DEC were just some of the eight inch floppy drives available on eBay in December 2013. All the technical manuals were available somewhere online and included detailed information on the power requirements and

interface cables too. Vintage computer clubs and hobbyists have made sources of documentation available for obscure media devices. The three authors visited the InfoAge Museum in Wall, New Jersey in 2012 and the people at the computer museum offered to help read any eight inch floppy diskette as well as help with any obsolete media needs.

When a thick eight inch by ten inch hard plastic diskette is encountered, there are other options such as using the IOMEGA Disk Cartridge System which contains a full length ISA card, cables, and eight inch Bernoulli Drive that can be connected to computers such as the HP Vectra that have room for ISA cards and contain USB ports. In figure 1 we can see a picture of obsolete media and various drives on a homemade system (HP Vectra) made from discarded computers and peripherals. The cover is off the unit so that more hardware is visible. Covers are very important to help prevent any shock hazards.

Figure 1 – Obsolete Media and Homemade Forensic PC

4 Dangers of Legacy Environments

In the 1990s, Dr. Doherty worked for Morris County ISD in Morris County, New Jersey and fixed computers, printers, and did a variety of data recovery work. There was an occasion where he was once called to fix a system that consisted of a PC connected to an IBM Quietwriter Printer. Dr. Doherty's tie was not clipped and got caught around the roller of the printer. His face was pulled closer until it lay on the printhead of the printer. The customer perceived a possible choking hazard and pulled the plug on the printer. Another danger that was potentially lethal was from the spinning rotisseries of the decollators getting caught on a person's tie. The decollator was used to separate three part paper into three separate collated and folded identical stacks of paper. This topic was discussed by Dick Brandon in a 1973 Computer Security Handbook [5]. Some devices have high voltage inputs and need to be handled carefully. Please consider devices such as the Shugart drive include three different voltages with the highest being about two hundred and twenty volts. Many PC examiners and hobbyists are used to working with a power supply whose output is five volts. Dick Brandon also discusses accidents from high voltage [5]. Such topics are seldom if ever mentioned in computer forensics or computer security academic materials in 2013 but in our opinion should be. This should be included in a MOOC or online class.

5. The Forensic Recovery Evidence Device (FRED)

Fairleigh Dickinson University (FDU) has a FRED Forensic Recovery Evidence Device. This device is made by the corporation known as Digital Intelligence and is used by many computer forensic investigators. There are many variations of the FRED such as the FRED Sr., FREDC, and FREDDIE [6]. The FRED Sr. can be used to examine a multitude of media types. A computer forensics adjunct who is now a retired computer crimes detective once showed a FRED station in his in person class at FDU. The value of such a device is discussed in chapter 6 of "Guide to Computer Forensics and Investigations" [6]. The FRED unit is also important because it has a write blocker and can be used to examine SCSI devices. It also contains a Firewire IDE interface with a write blocker. It is important to stress that such devices can be used to collect evidence and preserve it. A short video of the capability of such a device should be included in an online class or MOOC.

6. General Legal Concepts

It is important that the students learn about concepts such as the Spoliation of Evidence, Fruit of The Poisonous Tree, The Frye Test, and the Daubert Standard. Any MOOC or online class should emphasize that evidence from an improperly conducted investigation can be suppressed if it was not protected from being tampered with or if pirated software was used to obtain the results. The Fruit of The Poisonous Tree is a legal concept in the United States that relates to evidence being suppressed if it was the result of improperly licensed or pirated software. FDU is located in the State of New Jersey and it is Dr. Doherty's opinion that students should read a legal paper on the topic of spoliation of evidence by Lawrence Berezin ESQ. This paper called, "The History of
Spoliation in the Courts of N.J.: TARTAGLIA v. UBS PAINEWEBBER INC. 961 A.2d 1167 (2008) 197 N.J. 81 [7]." The paper is an important topic and may help emphasize the need to preserve evidence, handle it properly, and make it available to the correct people in the discovery process.

The Frye Test is important because students need to make sure that they are using proper techniques accepted by academics, the legal community, and law enforcement. Students should have a test plan of how they are going to

identify, collect, preserve, analyze, and then report the findings of digital evidence. The test plan should use practices and tools accepted by the previously mentioned communities so that it can pass the Frye Test in court [8]. It would be tragic if the results of an investigation were suppressed because the tools and methodology of the investigation were considered to be pseudo science.

7. CCTV Systems (Old & New)

Dr. Doherty created some informal classes about CCTV systems for visiting Cybercrime students and members of the Digital Forensics Special Interest Group (DFSIG) which meets once a week at FDU. Students appear to be most interested in CCTV systems and Digital Video Recorders (DVR). Most are interested in connecting the DVR to their LAN and viewing the video on their smartphone while away from home. IP cameras work well with a DVR and many systems are available online for approximately one hundred dollars. Many of the IP cameras include infrared technology and allow for night viewing. Since the in person DFSIG includes people of all ages who are interested in both new and legacy systems, some adapter/conversion equipment is introduced that allows the inclusion of old and new equipment on a CCTV/DVR system.

Some students also do some work in their community and have customers who are senior citizens with legacy security systems. Some continuing education students also work in the security field and occasionally say that they need help examining the footage from a CCTV system. The amazing thing is that sometimes the footage is from a legacy system that includes one camera and a VCR that uses a VHS tape. In 2013, there were still examples of 21 inch CCTV cameras from the 1970s and VCRs being sold on eBay. Many of these systems have a long life. These 1970 analog cameras could possibly be connected to a VCR which came on the consumer market in the 1970s. In the Eastern part of the United States, VHS was a dominant format. The problem with VHS tapes is that people use them too many times and that the resolution of the archived video is poor. Many people also use the cheapest tapes possible instead of more expensive high quality VHS tapes. This also lowers the quality of the archived video. Many people with VHS security systems do not use a cleaning tape to perform regular maintenance which also degrades the quality of the saved camera footage. There were many examples of vintage CCTV systems being advertised for sale on eBay which perpetuates the use of old systems. There are many companies that will transfer the entire VHS videotapes to a digital format on a DVD for a reasonable fee. Once the film is digitized, there are a variety of tools that can be used to magnify and enhance the video.

Students have asked how long CCTV systems have been around. They want to know the history for both their own curiosity and in case they are asked about the history of CCTV during the Voir Dire process in court. CCTV systems have been used since World War Two when German Rocket scientists preferred to watch rockets launch from a safe distance [10]. The German CCTV systems were brought to America along with German Rocket scientists and the V2 rockets in a special operation known as "Operation Paperclip." The InfoAge Museum in Wall, New Jersey has special archives and technology exhibits related to Operation Paperclip. Many of the captured scientists were given new identities and went with the CCTV and V2 rockets to the western United States where they worked on America's ballistic missile programs. The CCTV was then used for watching the results of the atomic testing. Students who visit Las Vegas may elect to stop at the Atomic Testing Museum on Flamingo Road to see some of the exhibits which also show CCTV cameras, drill heads, and test results. Since so much of the population travels, mentioning places such as the Atomic Testing Museum, an affiliate of the Smithsonian Museum, and the InfoAge Museum in New Jersey, offer people an opportunity to combine a leisure trip with learning. It is also worth noting that the CISSP is one of the premier network security certifications that computer science students and network security practitioners wish to obtain. A section of the CISSP exam contains some questions on CCTV cameras and CCTV systems.

8. File Systems and Formats

One of the most important topics in teaching legacy device forensics is data formats and file systems. Students should be taught to identify the type of device or media they have. They should find a brand, a model number, and look for a date. Then they can do an online search and find a picture and description of an identical device. The next step is to locate an online manual and find out the technical specifications of that device. The manual often discusses any format standards for the media or device. With media such as the eight inch disk, we may learn that it uses a FAT 12 file system. From there students can learn about file allocation tables, file names, clusters, pointers, and how files are added and deleted to the disk. This can help them recover files. Teaching students to use a disk editor can also help them look for data hidden near the MBR, master boot record, and in hidden partitions. Students should be encouraged to take a class on file systems with a local computer science department for obtaining deeper knowledge. Some eight inch diskettes such as the Memorex floppy disk used hard sectors instead of soft sectors. This gives the instructor an opportunity to teach the student about concepts in file systems that are rare but may occasionally cross an examiner's path somewhere in their career.

9. Conclusion

It appears that there is a pattern for teaching continuing education on legacy device and media forensics. There should be a section about identifying, preserving, analyzing, and reporting what was found and what it means. Identification includes finding a model number, date, data format, and instruction manual. Then one can understand what format the data is in and how it is retrieved and deleted. That can help one in selecting relevant tools for recovery and examination. The chain of custody is an important topic because the integrity of the seized media and device needs to be documented from the time it was taken until the moment it appears as evidence in court. Students should be taught about the history of the device and be prepared for the voir dire process where they are determined to be experts in court. The Frye Test and the need for using scientific methods accepted by the legal, academic, and practitioner communities needs to be stressed. The Fourth Amendment and search and seizure concepts should be stressed. Concepts such as the Fruit of the Poisonous Tree should be emphasized so that students do not use pirated software or stolen equipment to investigate legacy devices or media. Everything that can be done in person can also be done online if detailed high quality video is included along with supporting text and suggestions for further reading. Lastly, students should be encouraged to talk to experts and see legacy systems in action at places such as the InfoAge Museum.

10. References

[1]. Bischoff, G. (Jan. 2012). Why Pagers Still Matter. *Urgent Communications*, Jan2012, Vol. 30 Issue 1, p54-54. 1p, ISSN 1945-4384

[2]. Doherty, E. (2012). Digital Forensics for Handheld Devices. CRC Press, p81. ISBN 978-1-4398-9877-2

[3]. Elder, B. (2012). Chip-Off and JTAG Analysis for Mobile Device Forensics. *Evidence Technology Magazine*, Vol. 10, No3. Pp 10-15

[4]. *Shugart*. OEM Manual. SA800-SA801 Diskette Storage Drive – Shugart, p. 18, 23. URL Accessed 11/12/2013. Retrieved from : http://maben.homeip.net/static/S100/shugart/disk/shugart%20SA800%20Floppy%20OEM%20Manual.pdf.

[5]. Brandon, D. (1973). Computer Security Handbook. P. 99. NY : Macmillian Publishing. ISBN 0-02-469910-1

[6]. Nelson, B. et. Al. (2004). Guide to Computer Forensics and Investigations, p240-2. ISBN 0-619-13120-9

[7]. Berezin, L.(2008). The History of Spoliation in the Courts of N.J.: TARTAGLIA v. UBS PAINEWEBBER INC. 961 A.2d 1167 (2008) 197 N.J. 81. URL Accessed 11/12/2013, Retrieved from : http://shapiroberezins.com/wp-content/uploads/2011/08/Spoliation-history-nj-supreme-court-case.pdf

[8]. Smith, F. C., and Bace, R. G. (2003). A Guide to Forensic Testimony, The Art and Practice of Presenting Testimony as an Expert Technical Witness. Boston, MA: Pearson Education Inc., Addison Wesley. ISBN 0-201-75279-4, p 407-8

[9]. *IsoBuster*. Partial List of CD, DVD, and Blu Ray Formats. URL Accessed 11/26/2013. Retrieved from : http://www.isobuster.com/isobuster.php

[10]. *CovertVideo.com*. History of CCTV. URL Accessed 10/29/2013. Retrieved from : www.covertvideo.com/History_of_CCTV.htm

SESSION

E-GOVERNMENT AND ISSUES OF INTEREST TO GOVERNMENT, RELEVANT AGENCIES, AND PUBLIC ADMINISTRATION

Chair(s)

TBA

Scrivania: Public services execution and Semantic Search

Riccardo Cognini, Damiano Falcioni, Marco Maccari, Alberto Polzonetti and Barbara Re
Computer Science Division
University of Camerino
Camerino (MC), 62032, ITALY
{name.surname}@unicam.it

***Abstract*— Public Administration sector in modern society is characterized by the need to support extremely complex processes in order to provide services to citizen. In this area, transparency is one of the most important requirement to improve the quality of provided public services. In this paper we present a tool supporting PAs collaboration and transparency, named Scrivania. It allows Public Administration employee to model and publish their services and then be guided during the Business Processes model execution. Instead, citizens, using Scrivania, can search and execute the provided services, tracing their execution and, in case of delay, observe the state it occurs. Citizens can also be helped by a semantic search in order to execute specific tasks in a process.**

Keywords—Public Administration, BPMN, Collaborative Modeling, Guided Execution, Semantic Categorization

I. INTRODUCTION

Public Administration (PA) sector in modern society is characterized by the need to support extremely complex processes in order to provide services to citizens. Complexity is due to the fact that services provisioning is commonly a collaborative activity shared among different, possibly many, PA offices. PA services are in general interrelated, so the enhancement of activities involved in a service delivery can generally impact on other activities. In this domain, transparency is one of the most important requirement to improve administration efficiency and citizens satisfaction [1]. With transparency we mean the ability for the PA to make citizens aware of the delivery process in terms of activities and people involved in its execution and governance, improving the citizen's perceived trust. A service supports transparency if for each task, citizen knows exactly how to perform it and in case of delay know where is the bottleneck.

In order to support such scenario we have designed and developed a tool named *Scrivania* that implement transparency into PA and improve citizen's satisfaction about the e-government services usage. In this paper we focalize on the *Scrivania* use cases and how they are implemented in the collaborative BP environment that allow PAs to collaborate during the modeling phase of a service. This environment provides a real time collaborative BPMN 2.0 modeler that PA managers can use in order to design and provide PA services. Citizen can execute this services using the *Scrivania* services execution environment that integrate a semantic categorization engine. The engine is based on the state of the art in the related areas. It has been trained using English and Italian corpus documents and make use of a specific domain ontology in order to allow the correct categorization of common PA data and documents. Using this categorization is possible to improve the search and the collaborative functionalities offered by *Scrivania*, applying a semantic stamp.

The rest of the paper is organized as follows. Section 2 presents a background in BP modeling. In Section 3 the Italian e-gov scenario is presented. Section 4 presents *Scrivania* use cases. Then, Section 5 presents the architecture of *Scrivania*, services collaborative functionalities and semantic search during the execution of services. Finally, in Section 6 some conclusions are reported.

II. BACKGROUND ON BP MODELING

We refer to a BP as "a collection of related and structured activities undertaken by one or more organizations in order to pursue some particular goal. Within an organization, a BP results in the provisioning of services or in the production of goods for internal or external stakeholders. Moreover BPs are often interrelated since the execution of a BP often results in the activation of related BPs within the same or other organizations" [2]. BPM supports BP experts providing methods, techniques, and software to model, implement, execute and optimize BPs which involve humans, software applications, documents and other sources of information [3]. Recent work has shown that BP modeling has been identified as a fundamental phase in BPM. The quality of BPs resulting from the BP modeling phase is critical for the success of an organization. However, modeling BPs is a time-consuming and error-prone activity. Techniques which can help organizations to implement high-quality BPs, and to increase process modeling efficiency, has become an highly attractive topic both for industries and for the academy. Certainly many different commercial tools have been developed to support BPM. Nevertheless for what concerns the modeling phase they mainly provide support for BP editing and syntactical analysis. Different classes of languages to express BPs have been investigated and defined. There are general purpose and standardized languages, such as the BPMN 2.0, the UML Activity diagram, or the Event-Driven Process Chain and many others. There are also more academic related languages, like the Yet Another Work-flow Language based on Petri Nets

that's the most prominent example. Among the listed languages there are several differences. These are related to the level of rigor, going from semi-formal, with a precise syntax and with semantic given in natural language, to formal languages for which the semantic is provided thanks to well founded mathematical theories.

In our work we refer to BPMN 2.0 [5], an Object Management Group (OMG) standard. This is certainly the most used language in practical context. The standard specifies three views: process, choreography and collaboration. The process view refers to private and public BP. Using a private BP model intra-organization Business Processes are represented. At the same time in public BP the interactions between a private Business Process and another BP or Participant is modeled. The collaboration diagram, is used in order to have a complete representation both of internal process as well as of the message exchange structure. BPMN 2.0 collaboration given an intuitive graphical notation reported in Figure 1. In particular, the following BPMN 2.0 collaboration elements are the most commonly used in this diagram specification.

- Pools and lanes are used to represent a participant or an organization involved in the BP, and they contain the private BP and related elements as reported in the following.
- Tasks are used to represent an action to perform that produces a result. Different types of Tasks exist, just to cite a few we refer to manual, service, human, etc. Tasks are graphically drawn as rectangles with rounded corners.
- Events are used to represent something that can happen. In particular, Start events represent the points in which the BP starts, intermediate events represent something that can happen during the BP execution, like time exceeding a deadline or the reception of a message, and End events are raised when the BP terminates. Different types of events can be introduced starting from this three main categories. Events are graphically drawn as circles.
- Gateways are used to manage the process flow on choices and parallel activities. Different types of gateways are available, the most used are exclusive and parallel. An exclusive gateway gives the possibility to describe choices in the BP and a single output path can be activated each time the gateway is reached. Parallel gateway have to wait all their input flows to start and then all the output paths are started in parallel. Gateways are graphically drawn as diamonds.

Finally, choreography defines the expected behavior between interacting participants. In particular, the following BPMN 2.0 choreography elements are the same of collaboration except for task that includes information about participants. Choreography Tasks are used to represent the communication between two or more participants. They are graphically drawn as rectangles with rounded corners. Inside choreography task there are participants; the sender is drawn as a white rectangles and the receiver is drawn as a gray rectangles. They can be drawn upward or below in any order. Besides pools and lanes are not considered in BPMN 2.0 choreography because participant information are included in tasks.

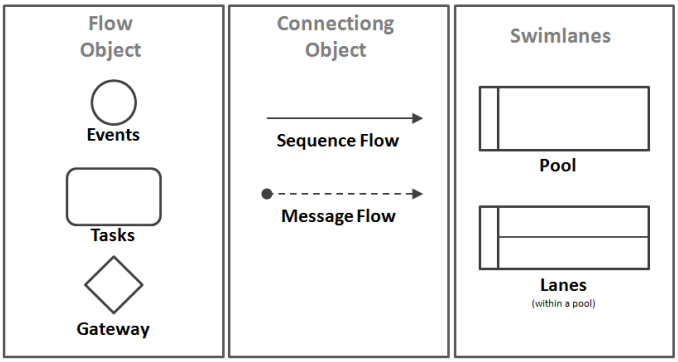


Figure 1. Example of a figure caption.

III. ITALIAN SCENARIO ANALYSIS AND THE APPROACH

The e-government in Italy is characterized by services availability but low service usage. This is common in small towns that represent 70.3% of all National Cities. From the point of view of the citizen there is an emerging demand for updated information and fully interactive public services. The citizens want to be aware about the to do list before using the service through the Internet. This is commonly done, asking friends or directly contacting the PA. This can all be traced in order to understand the level of citizen satisfaction and consequently review the administration process.

Moreover, e-government services cannot be based only on a single PA, but there are scenario characterized by complex inter-administration processes, where each PA contributes in relation of its responsibilities. In such a way, Public Administration communities can emerge, following the success story of popular social networks.

The proposed approach aims to improve the experience of citizens interacting with PA. With Scrivania we want ensure to PA a tool for continuous improvement.

The proposed scenario raises multidisciplinary issues that can be solved at technological level using Knowledge Management (KM) [6] and Business Process Management (BPM) [4]. Talk about KM and BPM in Public Administration means to have knowledge-intensive and collaborative processes. Scrivania address KM and BPM in two main areas: respectively (i) the citizens profiling and (ii) the service modeling.

First, in order to return to the citizen information and services more attractive than their own interests, Scrivania allow an accurate profiling of the individual. It can be implemented either with explicit and implicit mechanisms. The explicit profiling is done via the interpretation of bookmarks or through analysis of the registration data. The implicit profiling is done to taking into account several indicators such as the time spent on a page and the browser history.

Studies in Public Administration underscore the ongoing transformation on the service life cycle. The traditional processes in the public sector, historically driven by internal goals, are changing in a outside provision. That's mainly thanks to the evolving of new technologies.

The conventional model of PA who works as a separate and distinct entity, managing its own knowledge disconnected from each other, is being transformed. The new scenario is characterized by network relations in whom governments need to cooperate with other governments, with non-profit organizations, businesses and citizens to deal with the new challenges of globalization. New actors and stakeholders enter in relationship with the government. In this context it is essential to support administrative processes. Summing up, the transformation of PA service models, can be contextualized with respect to the following points as key factors of success: collaboration, control, sharing, transparency, inclusiveness and simplification.

IV. ACTORS AND USE CASES

The main actors of *Scrivania* are PAs employees and citizens. The most interesting use case related these actors are modeled using UML 2.0 as Figure 2 and 3 show. They provide a graphical representation of actor's interaction with *Scrivania*. Use cases can be divided in two groups:

- *Scrivania* services use cases are related to services creation and execution;
- *Scrivania* social environment use cases that are related to citizens social activities and interaction between actors.

In particular, *Scrivania* services use cases refer both to citizens and PA employs playing the role of administrator. As Figure 2 shows, citizens can search services, execute a selected service, check its state and evaluate the quality of the provided service. Instead, PA employs can create new services and make them available to citizens, manage services (for example in the case better versions of services are available) and check statistical indexes about services and citizen's opinions about tool and provided services.

Scrivania social environment use cases refer mainly to citizen as Figure 3 shows. They can search friends and send friendship requests. In this way they can communicate during a service execution if they want. Citizen can comment an executed service or an executed activity in a service, they can also read comments of friends related to services or activity. PAs and citizens employs a can be driven during service execution reading other citizen opinions.

V. THE TOOL

A. Overview on Scrivania architecture

Scrivania is implemented as a web based application running on Apache Tomcat. It permits to use java code directly to dynamic web pages. Thanks to jQuery we have implemented most of supported functionalities, for example the collaboration environment is created using jQuery post functions and java code. The used Database Management System is MySQL. It contains all the needed information such as citizens and administrators data, services models, services instances, etc. Then, *Scrivania* can be installed in each operative system because all the used technologies are free and open source.

Being a web application, PAs and citizens can use *Scrivania* via web browser (Internet Explorer, Firefox, Safari, etc).It means that *Scrivania* is a 3-tier application where the complexity of client is managed in the server.

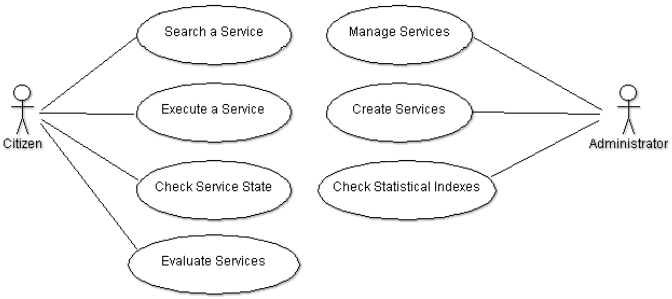


Figure 2. *Scrivania* Services Use Cases.

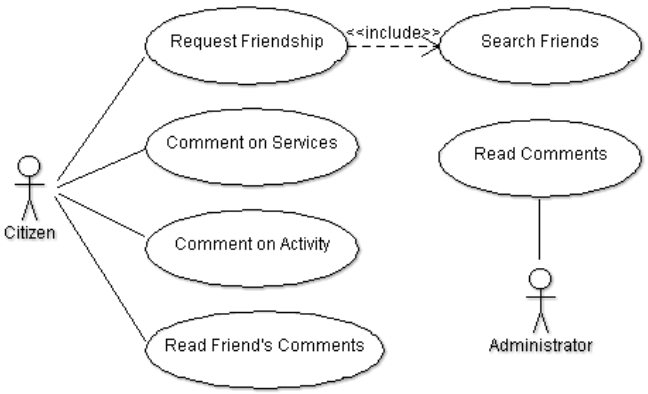


Figure 3. *Scrivania* Social Environment Use Cases.

The *Scrivania* architecture is based on a web application, suitable to support scalability and modularity. The main components are showed in Figure 4 and following described:

- *Authentication Module*, It allows citizens and administrators access to the functionalities provided by *Scrivania*. It guaranties three different levels of security that are following described. *Level 0* (self-registration) where users can register themselves and use such credential with a low level of security. *Level 1* (user name and password) is the most common and simple authentication system to administrate. It offers a lot of advantages, for example it does not need special hardware devices but it also presents the disadvantage that the association between the identity of people and authentication data is not guaranteed. Typically, this method is used to trace the activity of the user (profiling) and it grants a low protection level services access. Finally, *Level 2* (smart cards) is an authentication system based on physical support that guarantee the association between real identity and

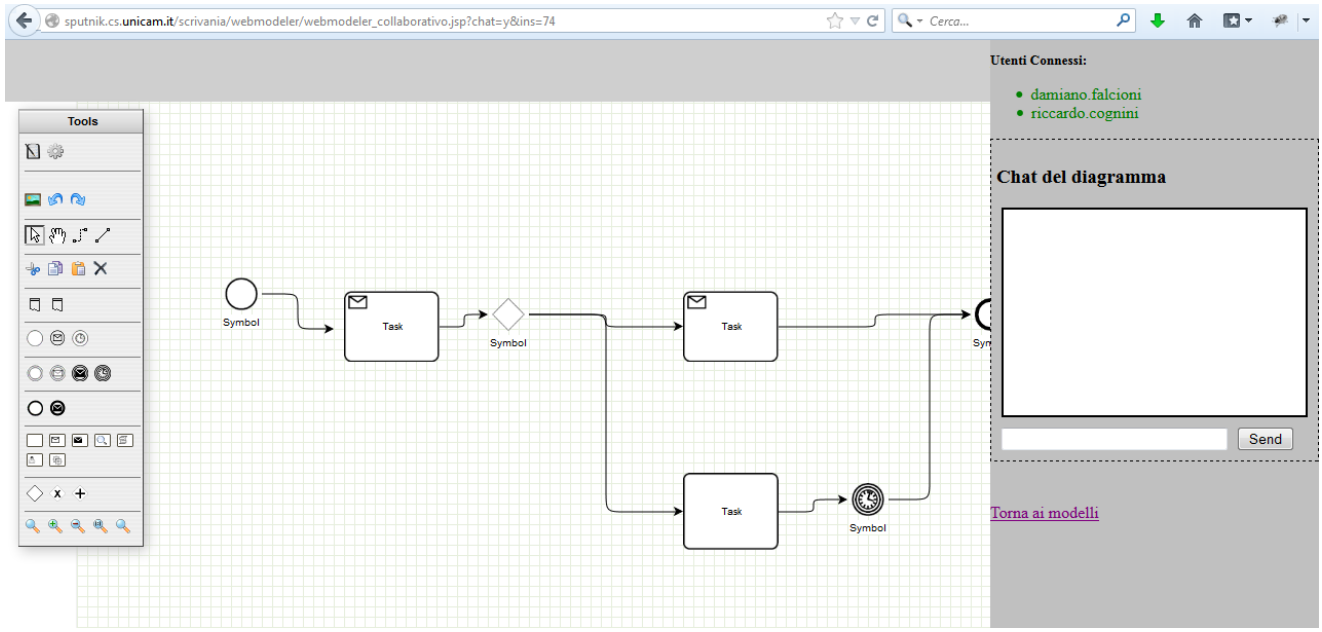


Figure 5. *Scrivania* collaborative BP environment – the BPMN 2.0 modeler

authentication data into smart cards. The security level can be further increase by a personal code that ensures the person from loss and robbery. In particular *Scrivania* supports the Electronic National identity card CNS as the national standard for digital identity card.

- *Administrator Module*, It allows administrators (PA manager) to manage PA services. Two different sub-modules compose this module that are Services Creation sub-module and Manage Services sub-module. The first provides functionalities for services creation; in particular it introduces a collaborative environment for BP modeling in which a BPMN 2.0 model can be designed. It enables communication between editors, during the modeling by a chat system. This sub-module support verification functionalities to guarantee the correctness of the model. Instead, Manage Services sub-module permits administrators to manage services and *Scrivania* functionalities. Using this sub-module, administrators can share services in order to allow citizens to their executions. Administrators can also check statistics about services quality.

- *Citizen Module*, It allows citizens to execute services and to interact each others. Three sub-modules compose this module. (1) Social sub-Module supports citizens communication during services executions. Citizen can also search their friend and give a feedback about a service or single activity of service.(2) Similarity sub-Module promotes services to citizen based on profile similarity. This means that probably two similar citizens need to execute same services. This sub-module provides an algorithm that merges two types of properties: service execution frequency and citizen similarity as a way of validation. (3) Execution sub-Module that provides functionalities to search and to execute services. It include also the semantic searching functionalities.

In this paper we focalize only on the collaborative functionalities of Services Creation sub-module and on the Execution sub-Module.

B. *Collaboration in the Services Creation Sub-Module*

Scrivania Services Creation sub-module provide a collaborative BP environment that PA managers use to create PA services in BPMN 2.0 standard. Each service has an owner that is the creator. The owner can enable other PA manager to edit the model during its creation. When the owner enable another PA manger to edit a model, *Scrivania* sends automatically a notification mail to the manager, containing the web link of the shared model. The environment is collaborative because each PA manager can see in real-time how other managers modify a specific PA service.

Figure 5 shows *Scrivania* collaborative BP environment. In the left we can see the toolbar of the modeler, instead in the right there is a list of active users, in this case damiano.falcioni and riccardo.cognini. Below the active user list there is a process chat used by PA manager to communicate.

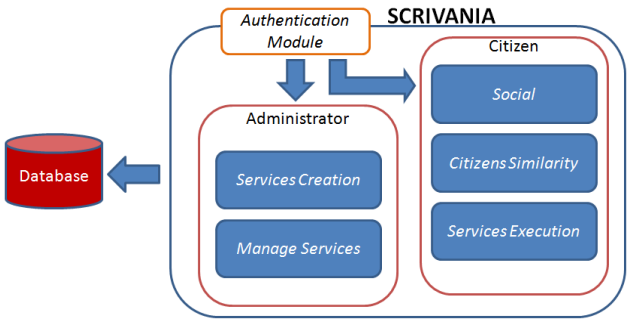


Figure 4. *Scrivania* Architecture.

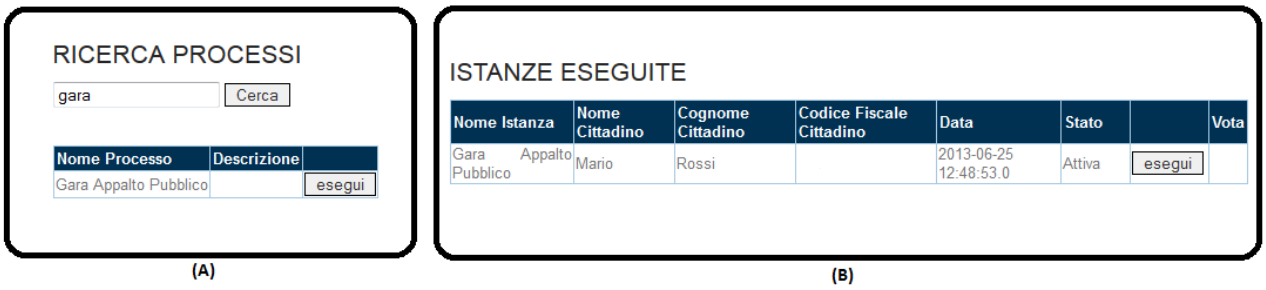


Figure 6. *Scrivania* services research (A) and citizen instance list (B)

C. Semantic Search of the Execution sub-Module

In order to execute a service, after its creation in the modeling phase, citizens can search it using the semantic search form as Figure 6-A shows. Search results shows the service name and provides a button to allow the process executions. When a citizen clicks on the execution button in his/her home page appear a new service instance as Figure 4-B shows. Each instance has a different instance name (that correspond to services name), a citizen instance, a date and the current status that can be active, in waiting and ended. An instance is active when the citizen still has to execute some tasks, while is in waiting state when he/she has to wait for other participant actions. An ended state means that the service has been completed correctly.

In an Active instance, when a citizen clicks on the execution button (button "esegui" in Figure 6-B) an new page shown in Figure 6-A appears, containing the list of all possible tasks to perform, as Figure 7-A shows. When the citizen choose to execute a task, a new page is opened, where there are some information about the selected task and the semantic search form of the semantic categorization engine as Figure 7-B shows. The search works on a semantic repository. The repository categorizes PA documents (or other kind of document if necessary) over a predefined domain ontology. This ontology has been defined in conjunction with different Italian PA domain expert in order to correctly represent all the information managed in the domain. The steps involved in the categorization of a document on the basis of ontological can be summarized as follows:

1. Text extraction and language identification of the document to be categorized;
2. Division of the text in sentences;
3. Part Of Speech Tagging of each sentence, in order to identify nouns, verbs, adjectives, articles, etc;
4. Dependency analysis to create the syntagms of the sentence and identify subjects, verbs and complements;
5. Comparison between relations defined in the domain ontology and syntagms identified.
6. Memorization of the resulting matches as metadata of the document, in a repository for subsequent search.

The semantic search is then performed as follows:

1. Language identification of the search query;
2. Part Of Speech Tagging of the sentence provided, in order to identify nouns, verbs, adjectives, articles, etc;
3. Dependency analysis to create the syntagms of the sentence and identify subjects, verbs and eventually complements;
4. Comparison between the syntagm identified and the relation stored in the repository and identification of the related documents to return.
5. Delivery of the documents to the user.

The framework used for the semantic features has been build from scratch, merging all the state of the art tools required for a semantic categorization of documents, taking into consideration adaptabilities feature to the Italian language, integration facility and performance. The Part Of Speech tagger and the Dependency Analyzer are in particular trained using a well affirmed open source Italian Corpus Document provided by the research group of the University of Trento.

VI. Conclusion

The paper presents a tool for Public Administration transparency named *Scrivania*. The tool provides to PAs an environment to model they services, in term of Business Process, in collaborative way. Citizen can search and execute the services they need using the execution environment that provides a semantic search in a documents repository in order to provide to citizen a high degree of transparency during the execution of services. Citizens can also use a social environment to communicate and helping each other.

Acknowledgment

We specially thank Consulthink s.r.l. partner of the project. The work reported in this paper has been founded by the Regional Tender "Avviso Pubblico POR FESR 2007/2013 – Co-Research, progetti di R&S in collaborazione, Regione Lazio (Italy)".

REFERENCES

[1] Bertot, John C. et al (2010) "Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies". Government Information Quarterly. Vol 27, No 3, pp 264-271.

[2] Lindsay, A. et al, (2003). Business Processes – attempts to find a definition. In Information & Software Technology, Vol. 45, No.15, pp 1015-1019.

[3] Weske, M. (2007). Business process management: concepts, methods, technology. Springer, Berlin, Germany.

[4] Weske, M. (2012). Business process management: concepts, languages, architectures. Springer.

[5] Object Management Group (OMG), 2011. Business Process Model And Notation (BPMN) Version 2.0

[6] Tiwana, A. (2000). The knowledge management toolkit: practical techniques for building a knowledge management system. Prentice Hall PTR.

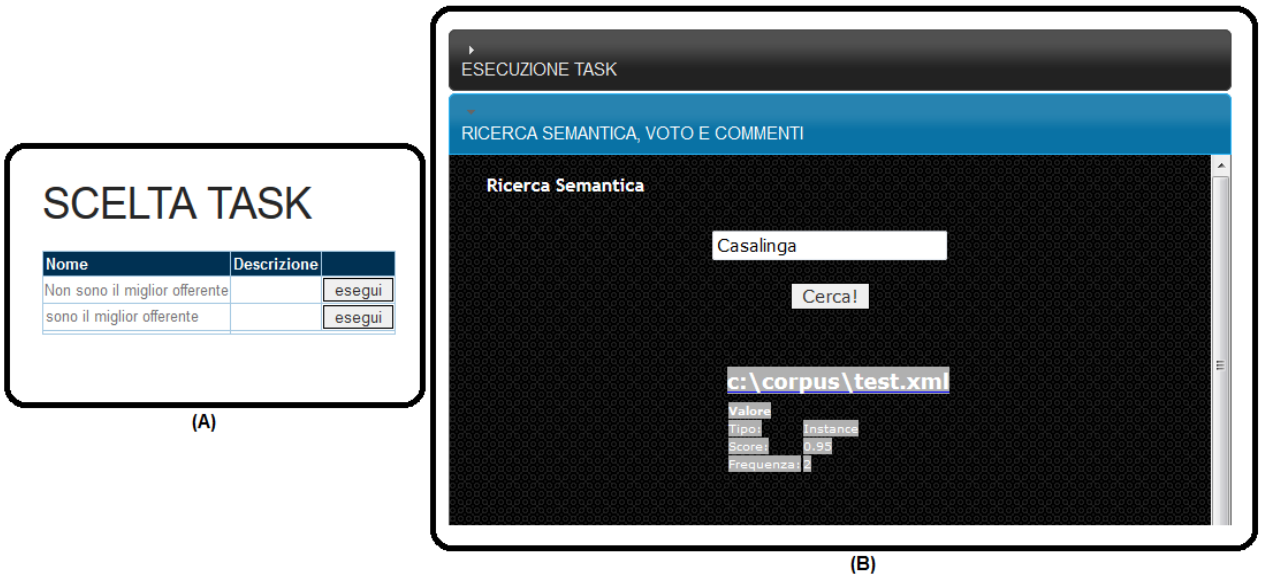


Figure 7. *Scrivania* list of tasks (A) and semantic search form (B)

Framework of e-government technical infrastructure. Case of Estonia

A. Kütt[1], **J. Priisalu**[1]
[1]Estonian Information System's Authority, Tallinn, Estonia

Abstract—*As information technology becomes more ubiquitous, governments of all levels must keep up with the increasing demand from the citizens for electronic services of all kinds. The service, policy and governance issues of this struggle are progressively well understood. Its technical aspects, however, are often considered to either be a commodity or covered by research in the field of enterprise architecture. Both experience in international cooperation attempts and efforts in coordinating architecture development on the state level, however, indicate that the existing approaches are not necessarily suitable for inter-enterprise coordination and governance purposes. In this paper we outline requirements for a framework suitable for framing the technical aspects of e-government, describe such a framework and discuss its application in the context of a small nation of Estonia.*

Keywords: e-government, framework, system architecture, software architecture, Estonia

1. Background and motivation

A country can be seen as a system performing a certain function by utilising a portfolio of limited resources. From this perspective, the concept of an architecture of a a country becomes meaningful. In the context of e-government, this architecture can be separated into elements dedicated to information technology and elements related to other aspects of the government. The latter is substantially covered by e-government related research while there is relatively little coverage of the former. In this paper, the authors seek to provide a conceptual architecture framework for the technical architecture of an e-government that is widely applicable and can act as a foundation of further research in the field.

2. Problem statement

Since the 1980s, there has been a solid stream of research focused on e-government, its structure, impact and governance. To a large extent, this work is focused on the service delivery and policy aspects of the issue while the technical implementation of the services and policies has received relatively little attention. Although architecture frameworks have been suggested (e.g. [1], [2]), they do not usually clearly separate between functional and technical aspects of information systems. Therefore application of these frameworks for technical architecture governance is limited.

In the following, we present three challenges driving the need for a structured understanding of the technical aspects of e-governance.

Firstly, e-government is often defined as the use of electronic means for service and data delivery ([3], [4]). While accurate in many ways, such an approach seems to be focused on existing services and assume the government, as a system, to have a stable architecture. This is not necessarily true as architecture developed by accretion might exhibit deficiencies in behaviour [5] and a definitive innovation pressure exists driving functional and technical change of which the open data movement is an excellent example. Therefore, stability of system architecture and the service portfolio, can not be assumed. Clearly, an understanding of the underlying technical architecture is necessary so the functional architecture it supports can be systematically developed.

Secondly there is evidence for system architecture being related to, and therefore influencing, organisational architecture ([6], [7]). A similar relationship has been observed in the field of knowledge management [8]. In the field of system architecture and product design, the notion of architecture having an unintentional impact on system functionality is known as emergent behaviour [5] or incidental interactions [9]. The research also points out, that not all such unforeseen behaviour is desirable. In the context of e-government, this means that decisions concerning technical architecture could have a potentially unwanted impact on the governance models in use and democracy in general. Understanding, predicting and ultimately controlling such effects assumes understanding of the architecture of the technical architecture of the government.

Thirdly, the governments are funded by tax-payers and are thus under varying levels of pressure to reduce costs. One of the most accessible cost-reduction mechanisms is consolidation of IT organisations as it can be achieved by administrative measures only. This however can damage the ability of organisations to fully utilise IT, a concept known as IT-business alignment. According to Luftman, IT-business alignment can be assessed based on six criteria: communications, competency/value measurements, governance, partnership, scope & architecture and skills [10]. At least two of these - communications and partnership - are

likely to be impacted as the organisational distance between IT and business is increased by centralising the former. Therefore, a need exists to identify more sophisticated means of consolidation which require a robust framework of thought around technical architecture.

3. Requirements

To be applicable and meet the challenges described above, the architecture framework to be developed should fulfil certain requirements.

1) It should provide a holistic view of the information system powering the governmental body in question in a way that is congruent with it in terms of scope, i.e. there should be one way of looking at the entire information system the said body is responsible for
2) Different governments have different approaches to performing their functions in terms of citizen involvement, centralisation, regulatory arrangements etc. A useful general framework should therefore fit a range of different forms of government while being flexible enough to accommodate for inevitable changes caused by democratic processes. It is also likely that the differences in approach (the concept part of the in the form-function-concept architecture model by Crawley [11]) can lead to significant differences between government architectures that should be possible to accommodate
3) The framework should have a level of abstraction that allows for conceptual discussion of the related topics. At the same time it needs to be specific enough to allow clearly define the related organisational architecture as a foundation for organisational change
4) As providers of e-government services vary in size and complexity from a small municipality to large confederations, the framework should allow for additional structure to be added to its elements in a way that does not affect the superstructure
5) All technical architectures are related to functional architectures, i.e. the way functional units are related to each other has an impact on how technical components interact. Thus, effective communication must exist between contributors to functional and technical architectures. In order to efficiently serve as a communication tool, the framework must, on a high level, be possible to describe in non-technical terms

4. The framework

Although the framework described is focused on technical architecture, technical architecture of any system is, as discussed, dependent on its functional architecture. Therefore, for each part of the framework, key question the functional architecture must answer are listed along with a brief description of their impact on architecture. These questions are centred around these three axis, each contributing to the complexity of the technical solution and the choices to be made:

- *Centralisation.* Generally a centralised solution offers more control but grows exponentially in terms of complexity for larger entities. A decentralised solution scales better but requires more complex coordination mechanisms to function
- *Privacy & security.* Any sharing of information or access to it causes privacy and security concerns. Thus, stricter policies in these fields drive up solution complexity while less restrictive policies require a framework for moving data between privacy domains
- *Diversity.* A more diverse solution space allows for a better match for potentially complex market needs and drives down complexity but requires robust coordination mechanisms. Stronger uniformity requires enforcement mechanisms to be in place and drives up complexity within a solution

The framework itself consists of four key layers uniting the information systems of various branches of the government and allowing interaction with the consumers of the services: citizens, officials and enterprises. Within each layer, the architecture of a particular government can be detailed using an enterprise architecture methodology of choice. The framework proposed intentionally does not cover the following aspects of e-government implementation:

- Implementation of the business logic of e-services themselves as the services are assumed to be implemented by the agencies and thus not to be subject to central governance
- Relationships between layers that, while undoubtedly existent, should be dealt with on a higher abstraction level

The framework is depicted on figure 1.

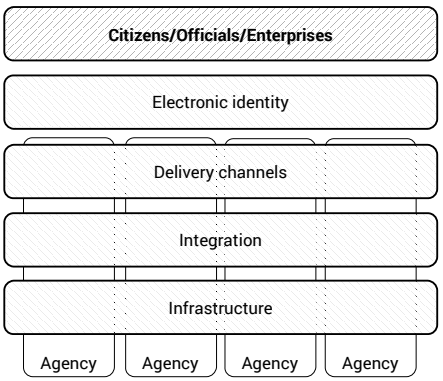


Fig. 1: Framework of e-government technical infrastructure.

4.1 Electronic identification

Without an efficient ubiquitous method of identifying a citizen and acquiring their legal consent, providing a full

range of electronic governmental services is difficult. At the same time, many services can be provided with either limited or no identification depending on both the risk levels present and the level of risk aversion of the service provider. The main functional drivers of the electronic identity layer implementation are:

1) Who is the target customer of the e-services provided? The question of how to identify a user depends heavily on who the user actually is. In EU context, countries are increasingly providing services to EU citizens in general. There are also circumstances where large portions of the population are not citizens of the country but might be considered a target group for e-services. The answer here is the foundation to many assumptions the technical implementation can make about the user behind the keyboard.
2) What is the legal significance of electronic identification? Identification methods carrying more legal significance require a more robust technical implementation and might be more difficult to distribute widely. Also, if the legal gravity of electronic identification is low, the services implemented must be designed to fill in the gap by, for example, requiring physically signing and returning a hard copy of the forms filled.
3) What is the multiplicity relationship between legal and electronic identities? Again, identification methods that provide a more certain relationship between a legal and electronic identity are more complex. Also, if a physical person can have multiple legally equivalent digital identities, a need arises to link these to each other via a shared identity code or some other mechanism. In case multiple identities can have various levels of legal significance and/or the identities can not be related to each other in a definitive manner, the information system implementing the architecture must make provisions to cater for the possibility of overlapping digital identities of various certainty levels.

4.2 Delivery channels

E-services can be provided via a range of electronic channels from traditional IVR-based ones to sophisticated mobile platforms. It must be noted, that the boundary between electronic and non-electronic delivery channels is not definitive as both kinds of channels support the same set of business processes and the systems implementing them can at least on the logical level be considered as one entity. The main functional questions driving the architecture are:

1) What is the diversity of the electronic delivery channels across services? In case a small number of channels is offered, accessibility issues arise as no electronic channel is truly ubiquitous. As the number of channels grows, the architectural complexity of keeping the service consistent across them increases.
2) What is the diversity of the electronic delivery channels across the country? It is often described as desirable to implement a "single-window" approach for providing services to the citizens ([12], [13]). This approach implies a system that needs to meet stringent availability requirements under high-throughput conditions and requires complex orchestration of systems of various agencies. A fully decentralised system, on the other extreme, would require mechanisms of providing a consistent user experience as well as a shared robust security domain.

4.3 Integration

The integration layer joins the information systems of different agencies allowing for sharing of data and functionality. This layer embodies the classical concept of middleware [14, p. 14] providing a clean separation between consumers and producers of data and services. In addition, the integration layer might provide assistive services like caching, service discovery, audit logging etc. The main questions driving the architecture of the integration layer are:

1) To what extent are the functions centralised between the agencies? This level of centralisation determines the ratio between producers and consumers of data and services. This in turn drives the availability, deployment and flexibility requirements towards the middleware platform.
2) What are the integration paradigms used? Paper-based governmental processes rely on documents being passed around. In the context of e-services, however, documents transform to data that can be shared, transferred and fragmented. In case of cross-agency processes, sharing data is often not sufficient as the agencies must coordinate different functionally significant activities and sharing of services becomes necessary. The choice between the three communication paradigms - document, data or service - has a strong impact on the requirements towards the integration layer. For example, a document-based system would need to provide tracking facilities for the documents moved while the service-based integration layer would need to consider transactional integrity.
3) How are the questions of privacy and data ownership treated? Since the main function of the integration layer is to allow access to data that often concerns citizens, the approach taken towards privacy is an important requirement driver for the architecture chosen as it is the main point of data access implementation. For example, the middleware might provide a facility for citizens to track access to data concerning them or enforce regulations preventing access to data or services under certain conditions.

4.4 Infrastructure

All of the services making up the e-government portfolio need to be deployed in a manner that supports the strategic goals of the e-government including non-functional requirements of performance, availability and security. The layer consists of the digital infrastructure - servers, networks, hosting facilities, security equipment etc.- used to provide e-government services. Its main functional drivers are:

1) How tightly is the infrastructure consolidated? Although there are definitive benefits like cost, high level of control and ease of integration to be gained from consolidating infrastructure, excessive consolidation can lead to emergence of single points of failure. Also, the extent to which the organisational structure of IT services is consolidated in the country have a strong effect here.
2) To what extent are platforms offered centrally? Emergence of cloud computing has shown the viability of centralised platform offerings. While the consolidation question can be seen acting on the horizontal dimension of the infrastructure layer, platforms can be seen as vertical. On one extreme of the scale is a slim platform that is limited to a small number of commoditised services like network access while sophisticated platform-as-a-service offerings are on the other.
3) What restrictions exist for the physical location of data? Both EU and USA have passed legislation that governs privacy of personal data and its transfer outside of the respective jurisdictions. In addition to these, countries might adopt additional restrictions while governments on the municipal level could be subject to a looser set of constraints or be required to share data with central government.

5. Framework application in Estonia

In the following the framework covered above is applied to describing the e-government architecture of Estonia, a small North-European country. In Estonia, this framework is used for two main purposes. Firstly, it forms the basis of communication within the country and with our partners abroad allowing for rapid identification of contact points and clean separation of concerns. Secondly, it is used as a governance tool with the national architecture governance process being aligned to the layers of the model.

For the electronic identification layer, Estonia is using a smart card that is also a compulsory picture ID from the age 15 onwards and is carried by vast majority of the population [15], including the residents without Estonian citizenship that account for 15.7% of the population [16]. The card is tied to a unique ID-code of a person. Authentication and digital signing are carried out using the certificates stored on the card, the resulting digital signature is legally equivalent to a physical one [17]. The legal significance of the ID-card is further elevated by its singular use for electronic voting in the country since 2005 and the rapid expansion of this voting method across several demographic dimensions [18]. Although other means of authentication, mainly federated authentication schemes provided by banks, are used in service offerings, they do not hold an equivalent legal weight. Estonia has also participated in an international effort for cross-use of electronic identification [19].

In terms of service delivery channels, Estonia has mainly focused on web based portals. In 2011, 94% of tax returns were filed electronically [15] via a traditional web-based service. More specifically, an internal unpublished analysis found 93,2% of the visitors to the main government portal in the first quarter of 2014 using a desktop computer with the rest using a mobile platform. This indicates a strong focus on traditional as opposed to mobile channels. Although a central government service portal exists, there are still more than 128 individual web-based service points as identified by an internal unpublished study. Thus it must be concluded the channels used in Estonia are not very centralised across the agencies.

For the integration layer, Estonia has chosen a relatively unconventional route and implemented a state-wide distributed integration bus called x-road [17]. X-road consists of a network of dispersedly deployed access points that mediate between agencies and the rest of the government infrastructure while performing support functions like service discovery and access control. This approach means all the data and services are distributed and all integration points are established peer to peer. Although x-road does not enforce a specific integration paradigm, majority of the APIs provided are data-based with service- and document based integration points being a clear minority. One of the main features of the platform is access control allowing fine-grained permissions to be set for data access. X-road also provides control points for data access monitoring.

6. Summary

We have outlined a framework that is designed to solve a number of challenges modern governments face. The framework is comprised of four layers - identification, channels, integration and infrastructure - providing a lattice uniting individual agency information systems into an unified system. This framework has been applied to a country of Estonia with two key positive outcomes:

1) The key aspects of the technical architecture of the country can be communicated both within the country and in the context of international cooperation
2) A holistic governance can be (and indeed is) applied to the technical aspect of the country allowing for systematic development of the services provided

References

[1] S. Sharma and J. N. Gupta, "Transforming to e-government: a framework," in *Proceedings of the 2nd European Conference on eGovernment-2002*. Academic Conferences Limited, 2002, p. 383.

[2] Z. Ebrahim and Z. Irani, "E-government adoption: architecture and barriers," *Business Process Management Journal*, vol. 11, no. 5, pp. 589–611, 2005.

[3] K. Layne and J. Lee, "Developing fully functional e-government: A four stage model," *Government information quarterly*, vol. 18, no. 2, pp. 122–136, 2001.

[4] Z. Fang, "E-government in digital era: concept, practice, and development," *International journal of the Computer, the Internet and management*, vol. 10, no. 2, pp. 1–22, 2002.

[5] E. Crawley, O. de Weck, S. Eppinger, C. Magee, J. Moses, W. Seering, J. Schindall, D. Wallace, and D. Whitney, "The influence of architecture in engineering systems," *Engineering Systems Monograph*, vol. 2006, 2004.

[6] A. MacCormack, C. Baldwin, and J. Rusnak, "Exploring the duality between product and organizational architectures: A test of the "mirroring" hypothesis," *Research Policy*, vol. 41, no. 8, pp. 1309–1324, 2012.

[7] J. E. Fountain, *Building the virtual state: Information technology and institutional change*. Brookings Institution Press, 2001.

[8] P. H. Hendriks and D. J. Vriens, "Knowledge-based systems and knowledge management: friends or foes?" *Information & Management*, vol. 35, no. 2, pp. 113–125, 1999.

[9] K. T. Ulrich, S. D. Eppinger, *et al.*, *Product design and development*. McGraw-Hill New York, 1995, vol. 384.

[10] J. Luftman, "Assessing business-it alignment maturity," *Strategies for information technology governance*, vol. 4, p. 99, 2004.

[11] E. Crawley. (2007, January) Introduction to System Architecture. Architecture to value. [Online]. Available: http://www.mitocw.espol.edu.ec/courses/engineering-systems-division/esd-34-system-architecture-january-iap-2007/lecture-notes/lec1.pdf

[12] A. Tat-Kei Ho, "Reinventing local governments and the e-government initiative," *Public administration review*, vol. 62, no. 4, pp. 434–444, 2002.

[13] J. Gant and D. Gant, "Web portals and their role in e-government," in *AMCIS 2001 proceedings*, 2001.

[14] P. Naur, B. Randell, F. L. Bauer, N. S. Committee, *et al.*, *Software Engineering: Report on a conference sponsored by the NATO SCIENCE COMMITTEE, Garmisch, Germany, 7th to 11th October 1968*. Scientific Affairs Division, NATO, 1969.

[15] Estonian ICT Export Cluster, "e-Estonia. The digital society," 2012.

[16] Ministry of Economic Affairs and Communications. (2014, March) About Estonia. Society. Citizenship http://estonia.eu/about-estonia/society/citizenship.html. [Online]. Available: http://estonia.eu/about-estonia/society/citizenship.html

[17] A. Kalja, A. Reitsakas, and N. Saard, "eGovernment in Estonia: Best practices," *Technology Management: A Unifying Discipline for Melting the Boundaries*, pp. 500–506, 2005.

[18] A. H. Trechsel, K. Vassil, *et al.*, "Internet voting in estonia-a comparative analysis of four elections since 2005. report for the council of europe," Council of Europe, Tech. Rep., 2010.

[19] A. Piñuela, S. Palomares, and A. Crespo, "D1.3.1 publishable summary report d1.3.1 publishable summary report d1.3.1 publishable summary report stork 2.0. publishable summary report," STORK 2.0 Consortium, Tech. Rep., 2013.

Income Inequality and Information Communications Technologies: Causes, Solutions, and One Example

Darren Gudmundson, MSc
Income-Inequality.com

Abstract - *Canada and the United States face growing in-country income inequality. Leading academic research from the International Monetary Fund (IMF), the Organization for Economic Cooperative Development (OECD), and other independent economists and political scientists have identified the contributing factors to increased in-country income inequality over the past thirty years as: changes to trade policy, changes to finance policy, changes to labour policy, changes to taxation policy, and, Information Communications Technologies (ICT).*

This paper articulates the economic environment in Canada and the United States and details the contributory role of ICT to the concentration of wealth and how ICT can be used to mitigate income inequality.

Keywords: Income Inequality, Android, Open Data

1 Introduction

Growing income inequality will be one of the greatest social and political challenges of our times [1].

From the earliest agrarian and hunter-gather societies from 7,000 B.C., through the Roman, Medieval, and Industrial periods, inequalities of income or resources have existed [2].

Given that income, or access to resources is based on individual physical attributes, intelligence, relationships, geography, birth right, and chance [3] - income inequality will likely continue as an integral part of the human experience.

However, over the last thirty years, in-country income inequality in Canada and the United States has grown significantly [4].

1.1 Measure of Income Inequality

The Gini coefficient is often used to examine and articulate income inequality. The Gini coefficient makes use of the Lorenz Curve. In economics, the Lorenz curve is used to examine distribution of resources among population groups, typically for the examination of income and income inequality. The Lorenz Curve is a representative two-dimensional plot of the ordered income of a population. The Gini coefficient is the size of the area between the actual distributive curve and the line of equality which is the 45 degree line from 0, which indicates perfect equal distribution to 1, which represents absolute consolidation of resources under the control of one individual, or entity [5]. It is the area between the Lorenz curve and the line of equality; the higher the co-efficient, the greater the inequality.

1.2 Canada - Income Inequality

When examined by the Gini coefficient, Market Income inequality in Canada for family units (includes individual households) has risen from 0.370 to 0.436, an increase of 17.8%. Looking at the disposable income (i.e.: after transfers) for the same population and time period, income inequality increased by 9.4%, from 0.286 to 0.313.

Market Income is defined by Statistics Canada as all income from all forms of employment, including the self-employed, as well as retirement income, and income from other sources. "After Tax", or "Disposable Income" is defined as all income after tax transfers and tax remittance [6].

1.3 United States - Income Inequality

The evidence for increasing in-country income inequality in the United States of America is evident. The United States Department of Commerce, Census Bureau reports a 18.4% increase in the Gini coefficient from 1980 to 2012 [7].

Based on the work of Piketty and Saez[8]; Luttig [9] notes that over the past thirty years, the distribution of income in the United States of America has shifted towards the very wealthy to the extent that it has become the most unequal of advanced democracies.

The Organisation for Economic Co-operation and Development (OECD) reported in the study, "*Divided We Stand: Why Inequality Keeps Rising*" that of the 34 OECD countries examined, when examining disposable income, the Gini for the United States in the late 2000s was surpassed only by Chile, Mexico, and Turkey [10].

1.4 Effects of Income Inequality

Income Inequality has serious implications for social outcomes like crime, poverty, life satisfaction, and health care [11].

Income inequality is a causal factor for depression, anxiety, reduced quality of life, and early morbidity rates for the elderly [12]. It is also a contributor to low birth weight, learning disabilities, and infant mortality [13].

With respect to a subjective general feeling of "well-being", income inequality within a nation is dependent on the norms and values of the citizens. In nations where the citizens place high value on trust and "doing the right thing" their feelings of well-being in times of high, or growing income inequality are less than those in nations where the citizens place a lesser value on trust and maintenance of a social

contract. The more egalitarian and trusting a society, the greater the damage is done to citizens by income inequality [14].

Income inequality is one of the casual factors in the 2011 "Arab Spring" that, through a series of often violent demonstrations, riots, and military actions; brought down the governments of Egypt, Tunisia, Libya and Yemen. In North America, particularly the United States, income inequality is a major driving force behind the "Occupy" and other socially disruptive movements [15].

While it would be facile to suggest that either Canada or the United States is on the verge of widespread civil disobedience and outright destruction of their political and economic structures, history makes it very clear that the lack of the "have-nots" to fully participate in society eventually brings about correction - or instability, and eventual destruction [16], [17], [18].

2 Causes

2.1 Trade

What is the role of trade with respect to growing income inequality in Canada and the United States and how has ICT enabled that?

The consensus among economists is that the liberalization of trade between nations increases the Gross Domestic Product (GDP) of the participating nations, and by extension the wealth of the people of the participating nations [19].

The consumer benefits from trade liberalization by having access a greater variety of products with increased quality at a lower cost. Business benefits by having access to markets they were previously excluded from, or had to bear the additional costs of tariffs and taxation as penalty for access to those markets. However, increased trade results in segments of the labour force being "displaced" and forced into lower paying jobs, long-term unemployment, or re-education [20].

Over the past thirty-three years, trade liberalization has been a priority for the governments of Canada and the United States. Since the middle of the nineteen-eighties, Canada has signed ten free trade agreements; and the United States, twenty.

The "rising tide lifts all boats" idiom stems from the works of legendary economists Adam Smith and David Ricardo [21]. This aphorism, attributed to John F. Kennedy is often used by proponents of supply side, or trickle-down economics while advocating for tax-cuts, reductions to government services, loosening of regulation, and increased trade with developing nations. To summarize, the argument is that if those that have wealth are allowed to do what they please with that wealth, then everyone, including the disadvantaged are better off [22]. Though the evidence shows that this belief may be mistaken [23].

Using an annual panel data set from 10 developed countries over 26 years (1980-2005) Cassette, et. al. [24] examined the short and long term impacts of trade on both goods and services. They conclude that over the short term, international trade has no impact on inequalities, but has a negative impact over the longer term. For goods, they conclude that there is a negative impact over both the short and long term.

In a 2004 study of the effect of the Internet on Trade, Caroline Freund and Diana Weinhold [25], the researchers used annual data (1995-1999) on the growth rates for exports and the number of Internet hosts per country to develop a bilateral trade model. Using the data and regression analysis they conclude that the Internet increases trade by lowering the fixed entry costs into new markets by providing nascent exporters and importers with relevant information at a low cost.

Without ICT it would not be possible to have the Global Production Networks that exist today. It is the Global Production Network that allows raw materials to be transformed in developing countries to finished or near finished goods for consumers in more developed nations. These supply chains are facilitated by trade liberalization, deregulation, and ICT [26].

2.2 Finance

In the context of growing income inequality and the financial crisis of 2008/2009, the area of finance; more specifically the financialization of economies of Canada and the United States, next to the enabling effects of ICT, financialization may be the largest contributor.

In classic economics the role of Finance is to provide a place for depositors to secure their financial assets; allocate credit across space and time; and, provide payment services. Strategically, the provision of these services enables organizations and individuals to hedge, price and share risk. Taken together, an efficient financial sector allows the producers to produce, the consumer to consume and overall provide the monetary liquidity that greases the wheels of a capitalist based economy [27].

Over the past thirty years the role of financial sector in Canada and the United States has grown dramatically, and changed considerably. It has transformed it's role from that of the subordinate provider of capital for the production of goods and provision of services, to one of, *"a predatory financial capitalism, where casino-type activities have become the norm — a process that is sometimes described as one of "financialization" with an almost complete decoupling of the financial sphere from commodity production.*" [28].

Driven by a neoliberal ideology of "the market knows best" and the movement from the Bretton Woods system where financial activities served the production and consumption actors in the economy, policy makers and regulators were bullied into backing off from their responsibilities to oversee the financial sector [29].

In addition to the spectacular collapse of the world financial system in 2008/2009, the principle impacts of financialization have been to: increase in-country income inequality and contribute to wage stagnation; give prominence and a priority position to the financial sector relative to the production sector; and, transfer income from the goods producing activities to the financial sector [30].

It is theorized that financialization accounts for as much as 70% of the growth of income inequality in the Anglo-Saxon system [31].

Financialization and the concentration of wealth that accompanied its rise would not have been possible without the explosive growth in the abilities and scope of ICT [32]. Regardless of changes to consumer preference, political structures, trade policy and labour practices, without the explosive growth of ICT, the financialization of the world economy would not have taken place [33].

2.3 Taxation

Even those most altruistic among us do not like paying taxes. However, outside of the most extreme libertarian schools of thought it is generally accepted that taxes are necessary to fund public services. Government expenditures on infrastructure, public safety, and national security are areas that consistently receive support [34].

In their paper, "The Rise of the Super-Rich: Power Resources, Taxes, Financial Markets, and the Dynamics of the Top 1 Percent, 1949 to 2008", Volscho & Kelly [35] draw upon a substantive body of social and political research to emphasize the connection between politics, policy, and income inequality. Their analysis, based on the Power Resources Theory[1] concludes that increasing income inequality (market and disposable) in the United States from 1980 to 2008 was due more too congressional shifts to the Republican Party; lower top tax rates; and, financial asset bubbles than due to market forces.

While arguably utilizing political influence may be enough to account for the shifting of the tax burden from the corporate to the individual, Boyd [36] notes that the shift of the burden of taxation from corporation to the individual has been in part due to multi-national corporations exploiting low tax rates in foreign countries. Wague [37] defines the Multinational Corporation (MNC) as having three characteristics: common control; common goals; and, common ownership of geographically spread resources–that create a paradox for nation states. The paradox being that while the MNC can bring foreign investment, the benefits are concentrated with the MNC and can be spirited out of the nation with the stroke of a key or click of a mouse. Further he observes that it has been the MNC that has been the primary facilitator and beneficiary of globalization enabled by ICT and government policy. In particular from the properties of ICT that enables the shifting of corporate resources from nation to nation to capitalize on differences between various tax regimes.

2.4 Labour

The decline of the role of Organized Labour in Canada and the United States is a contributing factor to growing income inequality [8].

Pontusson [38] found that from 1975 to 1995 countries where Unions fell out of favour and subsequently lost members and decreased in size relative to the workforce also suffered from rising income inequality.

This is further evidenced by Bruce Western and Jake Rosenfeld in their 2011 review of the role of Organized Labour and its impact on income inequality, "*Unions, Norms, and the Rise in U.S. Wage Inequality*" [39]. Western and Rosenfeld note that for the private sector, union membership dropped from 34 percent to 8 for men and from 16 to 6 percent for women for the period from 1973 to 2007. During that same time period, wage inequality in the private sector increased by 40 percent. Additionally, the researchers conclude that Unions reduce income inequality not only for their members, but for the workforce in general.

Without ICT enabling the global supply chain, the production of products in lower wage countries and subsequent export to Canadian and American markets would not be possible [40].

Kristal [41] defines "capitalists" as those who direct and own the means of production and "workers" as those who generate the wealth but are excluded from such ownership. She argues that "computerization" has contributed to the decline of organized labour, and by extension contributes to aggregate income inequality. This contribution is made in three ways; first through the elimination of unionized manufacturing jobs; second, union-busting or other anti-union actions; and finally, the polarization of skills in the workforce that diminishes worker solidarity.

Sachs [42] observes that labour laws that support the formulation of collective bargaining benefit the workers and reduce income inequality. However, he also notes that public policy in the United States in particular is far more responsive to the wealthy because their financial resources are put to work electing public officials that will do their beckoning.

Using yearly observations for each industry from 1967 to 1991; 1988 to 2007; 1978 to 2002; and, 1978 to 2002, Kristal [41] measured Labour's share of value by dividing Labour income by industry's value added. Her results conclude that even though arguments that increased incomes can be linked to biased skill-based tech change, when taken as an aggregate of the working population, the computerization of the workplace contributes to rising income inequality. She concludes that this is due to the fact that the investments in computerization are defacto investment in capital. That being the case, capital investment benefits the capitalist ahead of the worker.

3 Theory

There is widespread agreement that this concentration of wealth and political power have come to realization as a result of changes to trade, finance, taxation, and labour policies; enabled by Information Communications Technologies (ICT). Many of the studies pin the bulk of the heavy lifting being done by ICT [10], [43].

Without the ability to collect, manipulate, and transmit information around the globe, all in near real-time; just-in-time supply, mass customization, off-shore manufacturing,

[1] *Power Resources Theory links* class-based political power with income distribution. It was developed to explain differences in welfare [35] and posits that welfare states are more sophisticated, egalitarian, and advanced in countries where left-leaning parties and labor unions are stronger.

debt securitization, and financialization would not be possible [40].

This being the case, the questions remain, "can, and if so, how can ICT be used to mitigate income inequality?"

Without an agent through which to act, information is inert. It is only upon the application of the information that value is derived, or impact enforced. The individual, using the properties of ICT to gather, manipulate and disseminate information can act as an activist to educate, organize, activate with the goal of swaying political or business leaders and interests.

Fundamentally there are two ways in which ICT can be used to address income inequality; as an individual, or as part of a group.

The difference between an individual using ICT to address income inequality through influence and that of the Activist is that the latter undertakes a determined effort to organize other individuals within the population to contact, vote, boycott, or otherwise support the issue or cause [44].

Working with British Secondary School Students, researchers Rogers and Wild [45] identified the properties and potential benefits of using information technology in the classroom. Their empirical study identified several benefits of providing students with the ICT tools and access to data in the construction and interpretation of common graphing routines. They concluded the following properties of ICT and potential benefits:

Properties of ICT which add value	**Benefits**
New ways of gathering data.	Decreased dependency on secondary information. Increased scope of data acquisition.
Reduced manual effort to obtain graphs	Lower skill. Time bonus.
Real-time reporting	Increased interactivity. Time bonus. Promotes data gathering. Promotes data analysis.
Reading accuracy. Automation of data storage. Reporting accuracy.	Increased accuracy. Higher quality of information. Time bonus.
Automation of calculation and analytical assistance.	Increase of available information. Accuracy of derived information. Time bonus.
New methods of exploring data.	Increased analytical ability.

Though using IT in a classroom environment with Secondary School students may be different than attempting to educate a portion of the population on a specific issue, the properties of ICT that give the individual in the population group the ability to gather data and examine it as information in their contextual environment is transferable. As are the time saving properties, tools for calculation and analysis and new methods for exploring data. The question of accuracy of information is a property that is not transferable from the classroom environment to the Internet as currently deployed and is dependent on a given situation.

The property of enabling the individual to gather and lend context to information is a benefit to the Activist as it is of benefit to the free-thinking individual. The reduced manual effort and "time bonus" afforded by ICT in the class room is transferable to the Activist as well. The "Real-time" reporting property in the classroom environment permits interactivity and encourages more data gathering activities as well and encourages the student to seek out more data. These benefits can serve an Activist well when gathering information for the packaging of information for dissemination; responding to information being disseminated by the target organization; identifying and organizing supporters, and, research and communications.

The processes and contexts to effect change by non-state actors have been elucidated by King, [46]; Lipsky, [47], Myer & Minkoff, [48], and Amenta, et. al., [49], and the macro and elemental components of the use of ICT by non-state actors to effect change were articulated by Garrett [50].

With this construct it is possible to identify the specific contributions that ICT make to political or social activism in general, and by extension the practical use of ICT to address income inequality.

The properties of ICT greatly increase the ability of the Activist to frame the issue to mobilize the concerned individual to impact the opportunity structures.

In the context of this paper, the opportunity structures identified by Garrett [50] are known and understood as that of a Westminster Parliamentary Democracy for Canada, and Constitutional Republic for the United States of America. Both are democracies where the citizens have the right participation, representation, the right to assembly and generalized freedom of speech.

In this environment the government has the lawful right and responsibility to regulate organizations. That includes businesses, not-for-profit entities, and other elected bodies as directed by their respective constitutions and federal enabling pieces of legislation.

The regulated Organizations provide goods and services to the Population within the laws as established by the governments under whose jurisdiction they fall. Both Canada and the United States operate in a mostly free market Anglo-American style of Capitalism underpinned by the classic supply and demand model.

The population acts as consumers of goods and services from the Organizations, which includes businesses; and, additionally, although in declining numbers - as voters that select their representatives to a variety of governance bodies.

4 One Example

Operating ostensibly in a free market environment where the laws of supply and demand largely determine the return on investment requires that the business provide a product or service of value. The value of a product or service can be greatly affected by the opinion of the consumer of the

product or service, and the opinion the consumer of the providing businesses.

While there is no universally agreed upon framework or theory for consumer persuasion [51], and the traditional AIDA (Attention-Interest-Desire-Action) model has received criticism for not accounting for the consumers emotional state [52], it can be stated with empirical certainty that the emotional state of the consumer is a determinant of the ultimate purchasing decision, and that the consumer remains susceptible to influence right up until the moment of purchase [53].

In the field-test conducted from November 1, 2013 to January 31, 2014 an Android application (Shopping TriCorder) that allowed the user to scan in a Universal Product Code (UPC) using their smart phone or tablet was made available on the Google Play Store. The application presented to the user a list of the more popular retailers from which to choose. The application then allowed the user to scan in a Universal Product Code (UPC) and get back basic product information.

The user was also allowed to view the Executives of the retailer and the manufacturer along with how much compensation the Executive member received in 2012, or 2011 depending on data availability. Additionally, the user was provided with the Net Sales of the Retailer and Manufacturer, as well as their Net Profit, and how much the average employee for each organization was paid.

The user was also given the opportunity to participate in an online survey designed to elicit the impact of having this information near the time of intended purchase.

There were 65 responses to the questionnaire and when provided with information about the contribution of their intended purchase to growing income inequality, 12.7% changed their purchase; far from a majority, but not an insignificant number.

5 Conclusions

Information Communications Technologies (ICT) are enablers. Their collective ability to gather, manipulate, store, and control information have been used by business to automate production, coordinate distribution, displace labour, manipulate finance, and influence public policy.

They have also provided an opportunity for millions to improve their ability to earn income through employment in ICT, or, in the very rare cases, make Billions.

However, as noted, they have also been used to concentrate wealth and increase in-country income inequality in Canada and the US.

Though the number of consumers who changed their mind about a purchase when given information near the time of purchase was only 12.7%, that number is not insignificant. Millions of dollars can be spent to increase sales by that number, or to retain 12% of existing consumers.

Though consumers have been reticent to use UPC scans in their day-to-day shopping, many in the retail industry and manufacturers of retail consumer products are preparing for and promoting the use of RFID tags in the near-term retail space with the intent of pushing information to the consumer as they shop [54]. These businesses need to account for the use of the technologies they adopt and install by interests other than their own.

Conversely, the activist would be well served to be aware of, and position their organization to capitalize on this forecasted trend.

The Marketing Mix has been successful because it creates value for both the business and the consumer [55]. To effect substantive change to a market, policy, or practices of a multi-national corporation a large sustained effort over a period of time is required. As noted by Akpoyomare, et. al. [56] participating in a boycott, or movement for social change requires that the participant recognizes and identifies with an issue, has a desire to participate, and needs to receive a benefit for the cost of their participation.

As proven in this study, when provided with information near the time of their intended purchase, some consumers change their behavior. Given the participant model provided by Akpoyomare, et. al. [56] and from what is known about establishing and building relationships through the Marketing Mix, extending the functionality of an application like Shopping TriCorder to allow users to track their purchasing habits over time and be rewarded with virtual "trinkets", score, or currency; and compete or share their habits with other like-minded people may encourage committed participation over a longer period of time.

Through the use of smartphones, open source data and applications by one student, in a compressed time frame without the assistance of, or coordination with large and sophisticated special interest activist groups has shown that in a relatively short period of time with little difficulty - ICT can be used to sway consumer behaviour.

To fully assess the impact of the use of this type of technology on a market to change the remuneration to a more equitable sharing of resources, a larger sustained effort is required.

When examined from the viewpoint of how to use ICT to mitigate income inequality, though this project has proven that it is possible - in addition to the injection of information into the mind of the consumer near the time of purchase in an effort to sway consumer behavior, there remain several broader areas of research worthy of attention that cross several academic disciplines.

One of the unique properties of software and digital information, or data, is that though there are costs associated with the initial production, the cost of replication is as infinitesimally small as to be considered free. For software, like data and many forms of copyrighted materials there are a myriad of intellectual property laws around the world designed to assist the originator, or legal "owner" in their effort to recover their investment and turn a profit.

In classic market theory there is a labour cost associated with the production of new units. Aside from the initial investment to produce the digital material, there is no appreciable additional cost for the production of new units, labour or even material. This being the case, and given that an increasing amount of wealth is being created by

software, data, and other copyright materials; and, that the cost of replication of these materials is negligible – as we collectively move more towards a knowledge based economy can the classic model of supply and demand which underpins the neo-liberal Anglo-American political economy of Canada and the United States maintain relevancy and adequately serve the creation and distribution of wealth when the labour component is all but eliminated for much of a products life?

6 Acknowledgments

Thank you to my friends at the Product Open Data initiative who provide the world with free and unrestricted access to vital information that enables people to research the product they buy and put into their bodies.

7 References

[1] Beddoes, Z. (2012) *For Richer, For Poorer*, The Economist, [Online], Available at: http://www.economist.com/node/21564414, Last accessed: December 28, 2013

[2] Bentley, A., et. al. (2012) *Community Differentiation and Kinship Among Europe's First Farmers*, Proceedings of the National Academy of Sciences of the United States of America, Vol 9, No. 24, pp 9326-9330

[3] Calsamiglia, C. (2009) *Decentralizing Equality of Opportunity*, International Economic Review, Vol. 50, No. 1, pp 273 -290

[4] Kopczuk, W., Saez, E. & Song, J. (2010) *Earnings Inequality and Mobility in the United States: Evidence from Social Security Data Since 1937*, Quarterly Journal of Economics, Vol. 125, No. 1, pp 91-128

[5] De Maio, F. (2007) *Income Inequality Measures*, Journal of Epidemiol Community Health, Issue 61, pp 849-852

[6] Statistics Canada (2013) *Table 202-0709: Gini coefficients of market, total and after-tax income of individuals, where each individual is represented by their adjusted household income, by economic family type*, Government of Canada - Statistics Canada, [Online], Available at: http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=2020709&paSer=&pattern=&stByVal=1&p1=1&p2=50&tabMode=dataTable&csid=, Last accessed: December 9, 2013

[7] Department of Commerce (2013) *Income Gini Ration for Households by Race of Householder, All Races: GINIALLRH*, United States Department of Commerce: Census Bureau, [Online], Available at: http://research.stlouisfed.org/fred2/series/GINIALLRH/downloaddata?cid=33001, Last accessed: December 12, 2013

[8] Pikkety, T. & Saez, E. (2003) *Income Inequality in the United States: 1913-1998*, The Quarterly Journal of Economics, Vol. 118, No. 1, ISSN: 0033-5533

[9] Luttig, M. (2013) *The Structure of Inequality and Americans' Attitudes Toward Redistribution*, Public Opinion Quarterly, Vol. 77, No.3 pp 811-821

[10] OECD (2011) *An Overview of Growing Income Inequalities in OECD Countries: Main Findings,* Organization for Economic Co-operative Development (OECD), p38, [Online], Available at: http://egov.formez.it/sites/all/files/OCSE%20-%20An%20Overview%20of%20Growing%20Income%20Inequalities%20in%20OECD%20Countries.pdf.pdf, Last accessed: December 26, 2013

[11] Conference Board of Canada (2014) *Canadian Income Inequality: Is Canada Becoming More Unequal?*, Conference Board of Canada, [Online], Available at: http://www.conferenceboard.ca/hcp/hot-topics/caninequality.aspx, Last accessed: January 12, 2014

[12] Edvinsson, S. (2012) *Do Unequal Societies Cause Death Among Elderly? A Study of the Health Effects of Inequality in Swedish Municipalities in 2006*, Global Health Action, Vol. 6, pp 1-9

[13] Mayer, S. & Sarin, A. (2004) *Some Mechanisms Linking Economic Inequality and Infant Mortality, Social Sciene and Medicine*, Vol. 60, No. 3, pp 439-455

[14] Rozer, J. & Kraaykamp, G. (2013), *Income Inequality and Subjective Well-Being: A Cross-National Study on the Conditional Effects of Individual and National Characteristics*, Social Indicators Research, Vol. 113, No. 3, pp 1009-1023

[15] Bechtel, G. (2014) *Does Globalization Mitigate Income Inequality?*, Journal of Data Science, Vol. 12, pp 187-215, ISSN: 1680-743X

[16] Sen, A. & Foster, J. (1973) *On Economic Inequality*, Oxford University Press

[17] Sachs, J. (2001) *The Strategic Significance of Global Inequality*, Washington Quarterly, Vol. 24, No. 3

[18] Esteban, J. & Ray, D. (2011) *Linking Conflict to Inequality and Polarization, American Economic Review*, Vol. 101, No. 4, pp 1345-1374

[19] Whaples, R. (2009) *The Policy Views of American Economic Association Members: The Results of a New Survey*, Economists Journal Watch, Vol. 6, No. 3

[20] Langenfeld, J. & Niberding, J. (2005) *The Benefits of Free Trade to US Consumers: Quantitative Confirmation of Theoretical Expectation*, Business Economics, Vol. 40, No. 3, pp 41-51

[21] Mankiw, G. (2008) *Beyond the Noise on Free Trade*, New York Times, [Online], Available at: http://www.nytimes.com/2008/03/16/business/16view.html, Last accessed: December 8, 2013

[22] Stiglitz, J.E. (2002) *Globalization and Its Discontents*, Norton & Company, ISBN: 0-393-05124-2, pp 78-88

[23] Andrews, D., Jencks, C. & Leigh, A. (2011) *Do Rising Top Incomes Lift All Boats?*, B.E. Journal of Economic Analysis & Policy: Contributions to Economic Analysis & Policy, Vol. 11, No. 1, pp 1-43

[24] Cassette, et. al. (2012) *Income Inequalities and International Trade in Goods and Services: Short-and Long-Run Evidence*, International Trade Journal, Vol. 26, No. 3, pp223-254

[25] Freund, C. & Weinhold, D. (2004) *The Effect of the Internet on International Trade*, Journal of International Economics, Vol. 62, No. 1, pp 171-189

[26] Conchada, M. (2012) *The Role of Global Production Networks (GPN) in Understanding the Impacts of the United States' and China's Macroeconomic Stimulus on the ASEAN Economies,* DLSU Business & Economics Review, Vol. 22, No. 1, pp 1-26

[27] Herring, R. & Santomero, A. (1996) *The of the Financial Sector in Economic Performance*, The Wharton School – University of Pennsylvania, [Online], Available at: http://fic.wharton.upenn.edu/fic/papers/95/9508.pdf, Last accessed: December 10, 2013

[28] Barager, F. & Seccareccia, M. (2008) *Financial Restructuring: Implications of Recent Canadian Macroeconomic Developments*, Studies in Political Economy, Vol. 82, pp 61-83

[29] Panico, C., Pinto, A. & Anyul, M. (2012) *Income Distribution and the Size of the Financial Sector*, Cambridge Journal of Economics, Vol. 36

[30] Palley, T. (2007) *Financialization: What It Is and Why It Matters*, The Levy Economics Institute, [Online], Available at: http://www.levyinstitute.org/pubs/wp_525.pdf, Last accessed December 12, 2013

[31] Jacoby, S. (2008) *Finance and Labor: Perspectives on Risk, Inequality, and Democracy*, Comparative Labor Law & Policy Journal, Vol. 30, No. 1, pp 17-65

[32] Hope, W. (2010) *Time, Communication, and Financial Collapse*, International Journal of Communication, Vol. 4, pp 649-669, ISSN: 1932-8036

[33] McMahon, P. (2004) *Money, Markets and Microelectronics: Building the Infrastructure for the Global Finance Sector*, Prometheus, Col. 22, No. 1

[34] Donahue, A., et. al. (2008) *Taxes, Time, and Support for Security*, Public Budgeting and Finance, Vol. 28, No. 2, pp 69-86

[35] Volscho, T. & Kelly, N. (2012) *The Rise of the Super-Rich: Power Resources, Taxes, Financial Markets, and the Dynamics of the Top 1 Percent, 1949 to 2008*, American Sociological Review, Vol. 77, No. 5, pp 679-699

[36] Boyd, T. (2005) *Tax Reform in a Global Economy: Shifting the Tax Burden*, Review of Business, Vol. 26, No. 2, pp 48-52, 0034-6454

[37] Wague, C. (2013) *From Tax Break to Tax Haven – Offshore Profit Shifting and Multinational Corporations Tax Avoidance Strategy*, International Journal of Business Strategy, Vol. 13, No. 4, pp 115-128

[38] Pontusson, J. (2013) *Unionization, Inequality and Redistribution*, British Journal of Industrial Relations, Vol. 51, No. 4, pp 797-825

[39] Western, B. & Rosenfeld, J. (2011) *Unions, Norms, and the Rise in U.S. Wage Inequality*, American Sociological Review, Vol. 76, No. 4, pp 513-537

[40] Janvier-James, A. (2012) *A New Introduction to Supply Chains and Supply Chain Management: Definitions and Theories Perspective*, International Business Research, Vol. 5, No. 1

[41] Kristal, T. (2013) *The Capitalist Machine: Computerization, Workers' Power, and the Decline in Labors' Share Within U.S. Industries*, American Sociological Review, Vol. 78, No. 3, pp 361-389

[42] Sachs, B. (2013) *The Unbundled Union: Politics Without Collective Bargaining*, Yale Law Journal, Vol. 123, No. 1, pp 148-207, ISSN: 00440094

[43] Jaumotte, F., Lall, S., & Papgeorgiou (2013) *Rising Income Inequality: Technology, or Trade and Financial Globalization?*, International Monetary Fund, Vol. 61, No. 2., pp 271 – 309

[44] Lichterman, P. (1998) *What Do Movements Mean? The Value of Participant-Observation*, Qualitative Sociology, Vol. 21, No. 4, pp 402-418, ISSN: 0162-0436

[45] Rogers, L. & Wild, P. (1996) *Data-Logging Effects on Practical Science*, Journal of Computer Assisted Learning, Vol. 12, No. 3

[46] King, B. (2008) *A Political Mediation Model of Corporate Response to Social Movement Activism*, Administrative Science Quarterly, Vol. 53

[47] Lipsky, M. (1968) *Protest as a Political Resource*, The American Political Science Review, Vol. 62, No. 4

[48] Myer, D. & Minkoff, D. (2004) *Conceptualizing Political Opportunity*, Social Forces, Vol. 82, No. 4

[49] Amenta, et. al. (2005) *Age for Leisure? Political Mediation and the Impact of the Pension Movement on U.S. Old-Age Policy*, American Sociological Review, Vol. 70, No. 3, pp 516-538

[50] Garrett, R. (2006) *Protest in an Information Society: A Review of Literature on Social Movements & New ICTs*, Information, Communication & Society, Vol. 9, No. 2

[51] Meyers-Levy, J. & Malaviya, P. (1999) *Consumer' Processing of Persuasive Advertisements: And Integrative Framework of Persuasion Theories*, Journal of Marketing, Vol. 63, No. 4, pp 45-60

[52] Cramphorn, S. (2006) *How To Use Advertising the Build Brands: In Search of the Philophsers' Stone*, International Journal of Market Research, Vol. 48, No. 3

[53] Ravaja, N., et. al. (2013) *Predicting Purchase Decision: The Role of Hemispheric Asymmetry Over the Frontal Cortex*, Journal of Neuroscience, Psychology & Economics, Vol. 6, No. 1, pp 1-13

[54] Industry Canada (2012) *Chapter 2 – Consumers and Changing Retail Markets,* Government of Canada – Industry Canada – Office of Consumer Affairs, [Online], Last accessed: February 15, 2014

[55] Kotler, P., & Armstrong, G. (2012) *Principles of Marketing - 14th ed*., Prentice Hall

[56] Akpoyomare, O., et. al. (2012) *Consumer Motivation For Participation in Boycotts*, Business and Management Review, Vol. 2, No 9

The welcome brazilian idea about the United Nations supporting the regulation of Internet in housing rights, the probable failure of this idea.

A. Joel Mana Gonçalves[1], and B. Rodrigo Stábile[2]
[1]Directorate of Services to the Citizen, Prodesp - Cia. de Proc. De Dados do Estado de São Paulo, Taboão da Serra, SP, Brazil
[2]Legal Advice Management, Prodesp - Cia. de Proc. De Dados do Estado de São Paulo, Taboão da Serra, SP, Brazil

Abstract *- This paper discusses the idea of creating a Brazilian Multilateral Governance Pact on Internet, argument defended before at the UN, and the probable ineffectiveness practice of this measure, especially against the spirit of freedom of the Internet*

Keywords: *Internet; regulation; government; international laws; fundamental rights*

1. Introduction

This paper discusses the idea of creating a Brazilian Multilateral Governance Pact on Internet argument raised by the President of Brazil, Dilma Rousseff at the UN.

In spite of that salutary idea in his candid essence, in fact, the practical experience of the World Wide Web, it is feared that the bias aggressive and libertarian Internet becomes the measure, if approved, completely harmless and without any chance of achieving its purpose.

Corroborating, the fear that the new International Treaty turn dead letter finds room to germinate when they think of other examples experienced by humanity, with the turns violent actions of governments, despite there since 1948, a Charter of Rights supported by the UN.

Anyway, or the freedom that the internet is elementary condition, or lack of observance of governments to the text of treaties already signed, which in sight, unfortunately, is the likely ineffectiveness of the measure.

It is what goes on to explain.

2. The discourse and its arguments.

In the speech of the President of Brazil, Dilma Rousseff, at the United Nations in September 2013, due to the worldwide reported spying operation in which she and Brazil figured as victims, due to leaked documents by British journalist Edward Snowden, sustained the Brazilian head of state that the UN should consider a milestone civil multilateral of governance of the internet, a protocol to which all member states should obey. In his words, so stood the President of Brazil:

"The United Nations should play a leading role in efforts to regulate the behavior of the states facing these technologies and the importance of the internet, this social network, to build democracy in the world.

Therefore, Brazil will present proposals for the establishment of a landmark civil multilateral governance and use of internet and measures to ensure effective protection of data that travels through it.

We need to establish a global network multilateral mechanism:

1 – From freedom of expression, individual privacy and human rights.

2 – The Democratic governance, multilateral and open, exercised with transparency, stimulating collective creation and participation of society, governments and the private sector.

3 – From the universality that ensures social and human development and building inclusive societies and non-discriminatory.

4 – From cultural diversity without imposing beliefs, customs and values.

5 – From net neutrality to cover only technical and ethical criteria, making impermissible restrictions on political, commercial, religious or any other nature".

Analyzing the context of time, it is clear that the proof of the truth of the facts narrated converges to a violation of principles protected by international law, especially the sovereignty, and the dignity of the human person, regardless of this being the head of state and government of Brazil .

Following, if the intrusion as ventilated, has also been given in the commercial discovery of sensitive information and strategic Brazilian companies, it is not unreasonable distress to the UN, behold, BRAZIL, as a member of the United Nations, and traditions, never think to resolve this impasse in another field than dialogue and diplomacy.

It would be, on this scenario, again ratified the teachings of the Italian philosopher of our time, Norberto Bobbio, who argued: "the fundamental problem in relation to human rights today is not so much to justify them, but to protect them. This is a problem not philosophical but political. "

It would be, on this scenario, ratified the teachings of the Italian philosopher of our time, Norberto Bobbio[1], who argued: "the fundamental problem in relation to human rights today is not so much to justify them, but the protects them. This is a problem not philosophical but political.

"But with all due respect to the Brazilian position, the question is what causes anxiety, "- will solve?" Sorry to conclude with the pessimists. We explain:

Initially, it should be noted teachings of renowned Brazilian lawyer Patricia Peck Pinheiro[2], digital law expert, about the evolution of technology in recent decades until ratification of the massive use of the internet. Look up:

"Just over forty years ago, the Internet was just a project, the term "globalization" was not coined and data transmission over fiber optics did not exist ... The everyday world summed up the legal papers, paperwork and deadlines. With the changes that have occurred since we entered the era of real-time, the virtual displacement of business, breaking paradigms. This new era brings transformation in various segments of society ... The law is also influenced by this new reality. The dynamics of the information age requires a deeper change in the very way the Right is exercised and thought of their practice ... It is important to understand that we live in a unique moment, both in terms of technological and economic spheres of society. The professional in any field, especially the law, has an obligation to be in tune with the changes taking place in society. We know that the birth of the Internet is one of the major factors responsible for this moment, but the fundamental, first and foremost, is to understand that these developments are not the result of a cold reality, only technological, divorced from the everyday world.

The Internet is more than a simple means of electronic communication, formed not only by a global network of computers, but mainly by a global network of individuals".

In the same sense as a buffer of information as a right to be seen by all, the doctrine of the prestigious professor at the University of São Paulo Celso Antonio Pacheco Fiorillo and Christiany Pegorari Conte[3]:

"Infeasible discuss the advent of the Information Society without putting in a prominent position the Internet as well as their own reflections on the legal reality of the community. The information and communication technologies, especially the Internet, have brought the need for a new look at old rights, such as: information, communication, freedom of expression and to privacy, as well as questions about the emergence of new goods require specific legal protection (as in the case of so-called Computer Security, which covers the integrity of the information posted on the worldwide web, the availability of access to and confidentiality of information).

The "Internet problem" came to be identified when the technology began to interfere with the peaceful social relations and its subsidiaries, as well as some possible practices socially unpleasant and unwanted, such as its use in the commission of offenses and creating new contacts that jeopardize goods that have not yet recognized its relevance for the right.

The law must adapt to the new reality, under penalty of losing his true role, namely disciplinary social relations and enforce standards of conduct. Thus, the binomial law and the Internet is not passing phenomenon. It is a reality still underused, but that should be studied in all its fields of legal sciences, in order to secure new rights and the enforcement of existing ... We live in a society marked by the so-called Digital Revolution. Concepts such as the Internet, global village, virtual space and eliminating borders mark the social reality of the twenty-first century. In this context of virtual reality, new relationships are consolidated at each instant, requiring this way of legal protection in order to ensure effectiveness and safety for such relationships".

However, the greatest difficulty of any regulation in the field of internet is changing its essence free! The French thinker living in USA, Dominique Cardon[4] faced the issue:

"It happens that, in the digital era, democracy has changed in appearance. The Internet allows not only communicate more, better and faster, she extends formidably public space and transforming the very nature of democracy.

The Internet would somehow be the natural result of the evolution of mass media, since it can associate text, sound and image in digital multimedia. But the notion that the big chains temporally information media, is overly simplistic ... Suffice dominate this young rebel media to perpetuate that economic models, cultural and political set over the twentieth century".

Plus, even with numerous rules adopted by the international community for the preservation of fundamental rights and guarantees of human dignity, since the ending of the Second World War the various societies of mankind are obstacles to ensure minimum rights in many cases.

The Universal Declaration of Human Rights UN (1948) wove thirty articles parameters robust copyright protection for the first time in broad international perspective, propagating the ideals of freedom and respect between individuals and between states and their nationals, raising the dignity of the level of human worth more consideration.

In the words of the ever revered Italian philosopher Norberto Bobbio[5], there was embryo of the "Age of Rights."

As we know, the Universal Declaration of Rights of the United Nations was raised soon after the horrors and atrocities experienced by moisture in World War II – especially the ideals of Nazi extermination of other people. However, unfortunately, not so the world left to live in the years that followed clashes with new people, wars and suffering, often due to other interests that stems from

political and / or in many cases encouraging armed conflict until the collapse of one of the litigants.

Not infrequently these conflicts were mere splitting of shares from the post war 1945 – latent cases experienced by Eastern Europe in the 90s. On other occasions were due to political circumstances, accustomed to religious interests, political and commercial, as are examples of the Middle East, or the two U.S. military interventions in Iraq, for example.

In fact many other wars experienced humanity, without the vast majority of countries do not think of efforts to promote peace and honor respect the Charter of the UN Declaration of Rights, which is meant for itself a relevant tool in the search of the peace, however, not necessarily an effective guarantee. The Current situation in Syria to serve as a warning!

Incidentally, the difference is not so subtle way. Again using the teachings of Bobbio[5] we have:

"The problem of the foundation of a right presents itself differently depending on whether to seek the grounds that it has a right or entitlement that would have. In the first case investigate the positive legal system, to which I belong as having rights and duties, if there is a valid norm that recognizes and what is this rule, in the second case, try to find good reasons to defend the legitimacy of the right in question and to convince as many people as possible to recognize it ... We assume that human rights are desirable things, i.e., ends that deserve to be persecuted and recognized ... finishing by finding the reason and compelling argument, to which no one can refuse to own membership".

Later Italian philosopher concludes:

"The further increase the powers of individuals; especially diminish the freedoms of the same individuals. These are two different legal situations so that the arguments used for the first not worth to defend the second. The two main arguments for introducing freedoms between fundamental rights: a) the irreducibility of past beliefs b) the belief that the freer the individual is, the more he can progress morally and also promote the material progress of society. Now, these two arguments, the first is irrelevant to justify the requirement of new powers, while the second showed it as historically false. Well, two fundamental rights, but currently a contradiction, cannot have one and the other one absolute basis, i.e. a foundation that makes a right and its opposite, both irresistible and undeniable. Incidentally, it is worth recalling that, historically, the illusion of absolute foundation of some established rights was an obstacle to the introduction of new rights, fully or partially incompatible with those. Just think of the obstacles to progress posed by the social legislation of the absolute foundation of natural law theory of property: almost secular opposition against the introduction of social rights was made on behalf of the absolute foundation of rights of freedom".

Bailing the doctrine of Bobbio, what you want is to clarify the fact that sooner or later, with the internet, its benefits and harms, sees it possible that the Charter of Rights and Multilateral Guarantees Digital, as either the Brazil (e-Treaty) may be taken over by the UN, but honestly, do not believe that this solves the problem, nor inhibit any recurrence.

About the first it is good to say that with the internet or without the internet procedures are traditional espionage, the Cold War itself and its actors will say. But, obviously, that new technologies have facilitated and very fundamentally the internet, modalities intrusion, whether between individuals, companies, or countries.

Now, what you want to Brazil, via the UN, ultimately would transform the core DNA of the Internet: freedom! But the essence free is virtually impossible to be regulated in their minutiae, because even if a country does not promote invasion / intrusion a cracker could do it, despite being criminally culpable conduct.

In this sense, even the teachings of Liliana Minardi Paesani, Brazilian international lawyer and professor, warns that[6]:

"The law is always conservative, compared with the dynamics of the Internet, whose ability of new facts almost impossible lawmaker track their steps. Even being conservative, right cannot be silent and should seek to do justice, overcoming and adapting to the free nature of the Internet, in an attempt to preserve the rights of citizens, their privacy and integrity, blaming the offenders, even if virtual ... It appears that the idea of community is highly compromised. Private sphere and the public sphere, hitherto distinct, tend to mix it up. And there is a trend to replace traditional regulation by a kind of self-regulation relieved of political will, but The multiple connected networks of production and services".

And follows, basting the precepts of the Right to freedom of the Internet:

"The information system has become articulated and complex and has won a place in society and the ever-increasing recognition in the constitutional laws of the biggest countries in the world. The extent of expression of thought is strengthening and limits – and could not be otherwise – in numerous constitutional".

Studies renowned Professor Dominique Cardon, Sociologist and researcher at the EHESS in analyzing public space extended by the internet, remember:

"In the age of paparazzi, one must constantly remember that, since the late nineteenth century, the prominence of journalistic photography has raised many questions and this discussion prompted, then building a normative base for legally protecting the privacy of citizens. At the same time it develops a press hungry for news, Samuel Warren and Louis Brandeis theorized in 1890, the "right to peace". Since the activities of anonymous are not necessarily in the public interest, the systems were constructed in order to protect its existence against an invasion more violent towards the public. The porosity between the space of sociability and public space created the risk of personal information exposed for all to see. Á "institutional surveillance" of the state and enterprises around which the debates were organized on the personal overlaps currently

"a interpersonal surveillance" of a new kind ... The barrier has not disappeared off; it weakened".

Occurs, however, that analyzing the election by Brazilian approach, it is true that does not exist otherwise effective control giving prestige to the "irreducible ethical minimum"[7] acceptable in concrete actions of nations, this institute which, if violated, would imply sanctions to be discussed and applied by the international community, in fact, these penalties already exist, such as trade embargoes, requests for explanations, surveys, etc..

On the free internet universe has never been so crystal clear teaching of Bobbio, because the challenge is even protect the rights, and we do not see a Multilateral Pact between countries to be able to do it.

However, ensure their existence (the e-Treaty) is as welcome, even in the era of knowledge (fueled by technology) to repeat the same problems accruing to mankind after the 1948 rights era: There is an International Charter of Rights, so often violated, which is not desired.

In good time to note that in Brazil there is a Federal Law to dealing with digital crimes (committed by electronic means), Law 12,737 – known as Carolina Dickmann Law – famous actress of soap operas.

Professor Spencer Toth Sydow, Master of Law from the University of São Paulo , has very insightful work on the overview of computer crimes (as he prefers to call it), where we can withdraw from their work placements[8]:

"New concepts came into existence with the technology. While most goods was formerly represented by atoms, today most of them are represented by bits. Atoms form a tangible substance, while the bits make up the language (intangible) used by the computer to compose files, programs and communication signals.

Assets consist of bits also come into existence be achieved, thanks to new technologies. Trade secrets, copyrights, cash, databases, among many other values come into existence in the immaterial form.

On our times, there are no worldwide normative demonstrating behavior rules, in the computer network. Bet for a long time that the networks and their participants would be able to regulate themselves, simply by his communicative action, but the reality was not so friendly.

The power (of information) given to the connoisseur of language and procedures meant that if they experienced abuse in the network, exceeding reasonable limits of good living, a repetition of the ideas of Machiavelli.

It is noteworthy that computer crime is a kind of special crime, perhaps the sport "white collar", because practiced by educated and privileged groups, and assume that, in the opinion of experts ones, mostly the leadership among the crimes in the world in a very short time.

Offenders virtual criminals are not stereotyped in real society: not attending classes necessarily low, in general, need not committing criminal offenses to survive".

3. Conclusion

In conclusion, sustaining ourselves in the teachings of a renowned jurist and philosopher Brazilian Migual Reale[9], there is a foundation for the idea of Brazilian President:

"Every human history binds primarily to the history of law, for whatever a man does, from the great artistic achievements to the humblest domestic uses, everything is conditioned, directly or indirectly, by law or manifests itself through legal forms. The right is not the highest value, but what determines the other, it is not life, but the fundamental warranty of society…

The law as cultural force it is, and, more precisely, as an element of order and ensuring the cultural values of a community, can not only be stable - that is stagnation and death - not only movement and change, which is the lack continuity, breakage, waste of life ...Every culture has a center's own values, or rather several "centers of value" side arranged around a core value because, in the succession of cultural cycles, a value is essential: the human person, we have already pointed as the source value".

About the universal source value of the rights on the most advanced societies and cultures, we emphasis the appointments of Professor Liliana Minardi Paesani:

"The means of mass communication, enhanced by new technologies, break cultural boundaries, political, religious and economic. Internationalization of information was anchored on the doctrinal movement that seeks to protect the interest of the individual goals and, in this case, the collective interest connected to the computer and telecommunication. The sector of mass communication today is one of those areas, and most say the new rights of the people in its essence Community.

The law is always conservative, compared with the dynamics of the Internet, whose ability of new facts almost impossible lawmaker track their steps. Even being conservative, right cannot be silent and should seek to do justice, overcoming and adapting to the free nature of the Internet, in an attempt to preserve the rights of citizens, their privacy and integrity, blaming the offenders, even if virtual

It appears that the idea of community is highly compromised. Private sphere and the public sphere, hitherto distinct, tend to mix it up. And there is a trend to replace traditional regulation by a kind of detached self-regulation of political will, but connected The multiple production networks and services

'The Power, when abused, ceases to be legitimate and degenerates into tyranny; Freedom, in turn, when taken to excess, ends up producing the same effects.' (Benjamin Franklin)

The evolution of State forms (State Absolute State Liberal) and the affirmation of the welfare state that recognized and placed first in the economic and social rights highlight the current need to include other rights.

Today, scholars fall into the rights and freedoms "generations" linked to the historical period in which it is stated by those documents that was its precedents.

Freedoms and rights of first generation, comprising all freedoms of individual character;

Freedoms and rights of the second generation, composed of economic, social and cultural rights;

Freedoms and rights of third generation, also known as solidarity rights not for the individual, but directed to the social group (people). It is the right of peoples to self-determination, peace, development, ecological balance, the control of national resources and environmental protection;

Freedoms and rights of the fourth generation, Rights are being recognized for the field of genetic engineering, bioethics and new communication technologies.

This convergence of thoughts in conclusion, no technology and their inventions are increasingly fascinating and intriguing, yet humanity is in the hands of one of their ancient dilemmas - free will, tied to ethics, education and culture of the people is that will determine what is acceptable (ethical minimum), and what is at stake!".

4. References

[1] Bobbio, Norberto, "1909-A era dos Direitos". Translation of Carlos Nelson Coutinho; apresentação de Celso Lafer. Nova editora – Rio de Janeiro: Elsevier, 2004. – 13ª reimpressão.

[2] Pinheiro, Patricia Peck, Direito Digital. 4 ed. Ver. Atual. E ampl. São Paulo. Saraiva. 2010. p 43.

[3] Fiorllo, Celso Antonio Pacheco – Conte, Christiany Pegorari - Crimes no Meio Ambiente Digital – São Paulo: Saraiva, 2013.

[4] Cardon, Dominique, 1965 – A democracia Internet: promessas e limites. Tradução Nina Vicente e Tiago Coutinho; Revisão técnica Manoel Barros da Motta. – Rio de Janeiro: Forense Universitária, 2012.

[5] Bobbio, Norberto, "1909-A era dos Direitos". Translation of Carlos Nelson Coutinho; apresentação de Celso Lafer. Nova editora – Rio de Janeiro: Elsevier, 2004. – 13ª reimpressão.pp 15 e seguintes.

[6] Paesani. Liliana Minardi, Direito e Internet: LIBERDADE DE INFORMAÇÃO, PRIVACIDADE E RESPONSABILIDADE CIVIL. 5ed. São Paulo. Atlas. 2012.

[7] Piovesan, Flavia. Direitos humanos e justiça internacional: Um estudo comparativo dos sistemas regionais europeu, interamericano e africano – 4 e. revisada e ampliada. São Paulo: Saraiva, 2013, p 45.

[8] Sydow, Spencer Toth. Crimes informáticos e suas vítimas/ Spencer Toth Sydow; coordenadores Alice Bianchini, Ivan Luís Marques e Luiz Flávio Gomes – São Paulo: Saraiva, 2013 – (Coleção saberes monigráficos)p. 06/07

[9] Reale, Miguel. "1910 – Horizontes do direitor e da história". 3ª Edição, verificada e aumentada. São Paulo: Saraiva, 1999, pp.284, 292-293.

E-GOVERNANCE FOR EFFICIENT MANAGEMENT TO REDUCE CORRUPTION: AN ICT DRIVEN PARADIGM

T.Sheela,Sri Sairam Engineering College, Anna University, Chennai, India
hod.it@sairam.edu.in
Daniel Chandran, Faculty of Engineering and IT, University of Technology Sydney, Australia
Daniel.Chandran@uts.edu.au

Abstract:

E-governance is the civil, political and economic processes of government including service provision, using information and communication technology. Only 25% of e-governance is completely successful in India as on March 2014.(Source: InformationWeek). The trail of unaccounted money (black money) is running round the clock in recent times. To curtail the flow of black money, illegal property purchase and proper tax levitation, a new routine called THE MONEY CARD SYSTEM is proposed in this paper. Purchase of any product or service has to be done only through this card by swiping it using a AAA (Authentication, Authorization, Accounting) for fund transfer. The money transfer takes place through e-banking. The database of the bank containing the information of the user will be put on cloud. SaaS (Software as a Service) provided by cloud can be used to compute the total possible income and expenditure for a person. Transparency can be achieved by providing an application in the cloud like Gmail to interface the user to account his/her details of income and expenditure. But in a practical platform using card is difficult. Hence every individual is allowed to hold some amount of threshold money (which is 10 % of a person's income) in their hand. The money cannot be withdrawn more than once a week and details has to be furnished regarding the expenditure of the threshold money which will be accounted in their account.

Keywords: E-Governance, Cloud Computing, Card System, Arthakranti ,Hadoop Service

INTRODUCTION:

Nearly all countries have integrated Information Technology into their national economic strategies to improve the quality of life of their people. E-government aims to provide better services to their citizens. OECD defines e-government as the use of information and communication technologies, and particularly the Internet, as a tool to achieve better government. There is a significant increase in e-government operation throughout the world. Developing countries are still behind in providing e-government services. One of the services that can be offered to citizens is Government to Citizens (G2C) which could be of benefit to the citizens and the government, such as online tax filing, personal documents like passports, birth certificates, property development applications, etc.

E-governance is the civil, political and economic processes of government including service provision, using information and communication technology. Governments, in particular, in the developing countries are in the domain of cloud computing, planning new ways to interact, improve services, optimise processes and revitalise democracy by spending their resources on IT. The main target is to deliver more interactive services to citizens and businesses through e-governance. For this purpose, the cloud computing helps in significant cost savings, wherein computing hardware and software infrastructure and applications are remotely hosted. The paper describes how this newly emerged paradigm of cloud computing can be helpful for e-governance. The existing e-governance is server centric, cost effective in nature and finds it unable to address all categories of users from rural urban to metropolitan citizens. Hence e-governance facilities are confined within limited inhabitants of India and remain unsuccessful. In this paper, we propose a new effective framework of e-governance based on cloud computing concept. One of the important sources of income to government is through tax collection. The present available system in inefficient in monitoring a proper tax collection and it is also a tedious process for the tax payers to pay their taxes. Hence we provide new routine, *THE MONEY CARD SYSTEM* to curtail the flow of unauthorised money (black money) and proper tax levitation. The System proposed has the tax monitoring capability with greater ease and comfort.

PROBLEM STATEMENT:

The current trend of e-governance, though it satisfies many problems, is not yet completely reliable and sustainable for increasing the efficiency, reducing the corruption and circulation of black money. The current basic problem is the system of tax collection. There is a need for an ideological as well as a structural change in the e-governance system. To govern in the context of government is to exercise the authority to control and direct the policies.

OVERVIEW OF E-GOVERNANCE IN INDIA

The process of electronically filing income tax returns through the Internet is known as e-filing. It is mandatory for companies and firms requiring Section 44AB audit to submit income tax returns electronically. Clients requiring Section 44AB audit return without an e-filing receipt will not be accepted. The Income Tax Department is keen to encourage e-filing of IT returns by all taxpayers in view of the following benefits to taxpayers.

- Anywhere-Anytime Filing
- No long queues
- Quick Processing
- Accurate data in return

The new Income Tax (IT) return forms were first introduced for AY 2008-09 in March 2008. The Income Tax Department is presently preparing their internal software to enable e-return preparation and to accept the e-filed return. IT return form scan be download.

E-filing process

The e-filing process is briefly described in Fig 1. Tax payers can use Online ITR Filing program or download STO – ITR Software. The **ITR Software of Saral Tax Office** makes the process simple. STO - ITR Software is linked with STO - Income Tax Software, Balance Sheet, etc to get the data directly from the clients. User can also directly make the entry and generate the Returns either on paper or in XML.

Manual Filing

In case of manual filing, individuals have to audit their income and pay their taxes accordingly. The forms like ITR1,ITR2,ITR3,ITR4,5 are filled depending on the nature of work.

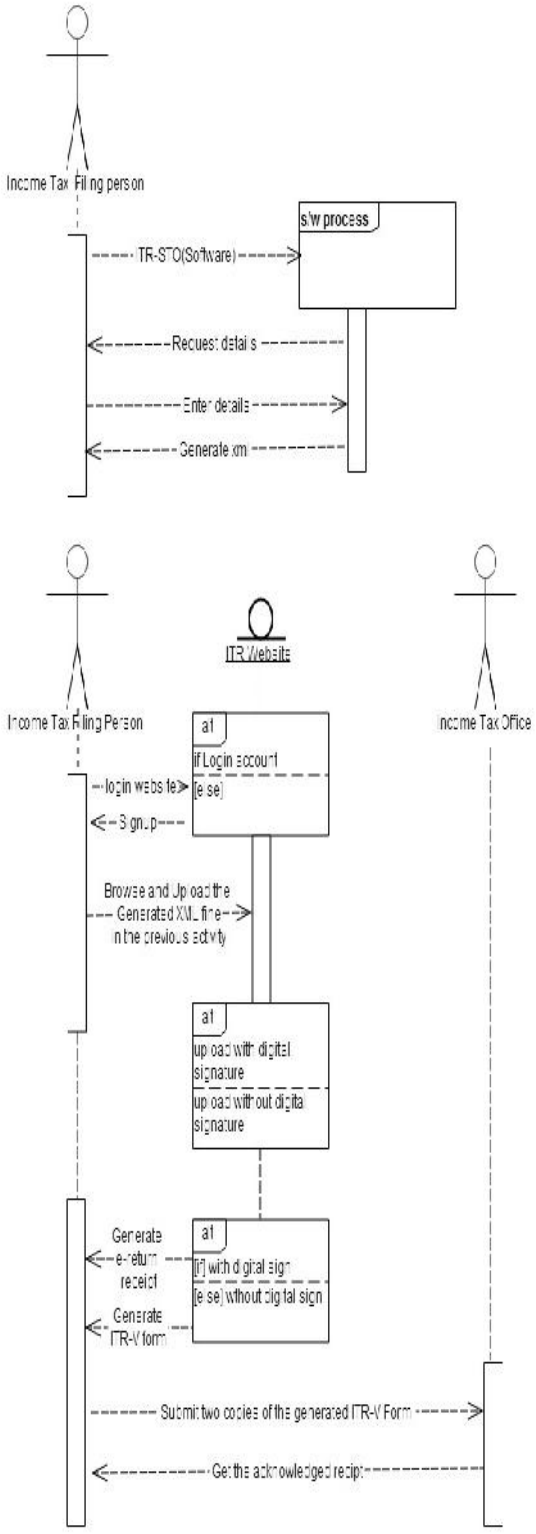


Fig.1. ITR filing process-UML diagram

Components of the present E-governance system.

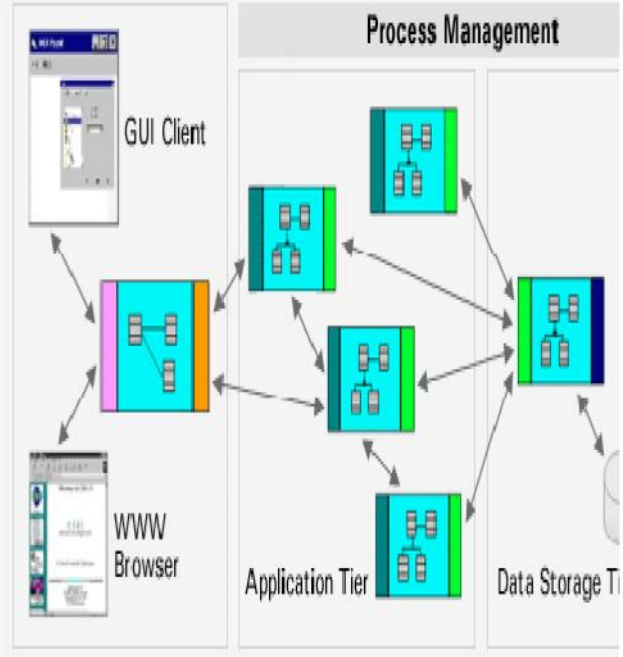


Fig. 2CurrentE-governance system

ARTHAKRANTI PROPOSAL

Arthakranti proposal has come out of Pune based Arthakranti Sansthan, an Economic Advisory body constituted by a group of Chartered Accountants and Engineers. This idea has been patented by the Sansthan. The Proposal aims at solving all socio-economic issues. "The solution consists of some simple technical corrections to our systems of economic governance" (ArthaKranti). It aims to provide an effective and guaranteed solution to black money generation, price rise and inflation, corruption, fiscal deficit, unemployment, ransom, GDP and industrial growth, terrorism and good governance.

Arthakranti Proposal has 5 points of action:

- Withdrawal of existing taxation system completely, except customs import duties.
- Transaction tax for every transaction routed through a bank will attract certain deduction in appropriate percentage. i.e. a single point tax deducted at source. All high value transaction has to be made through a banking system.
- Withdrawal of high denomination currency from circulation, e.g. 500 and 1000 rupee notes.
- Cash transactions will not attract any transaction tax
- Government to make legal provisions to restrict cash transactions up to a certain limit e.g. Rs.2000. (ArthaKranti).

Current status:

1. The total banking transaction is more than 2.7Trillionrupeesper day which is, more than 800 Trillion rupees annually.
2. Less than 20% transaction is made through banking system and more than 80% transaction is made in cash which is not traceable.
3. 78% of Indian Population spend around 20-100 rupees daily. Use of 1000 rupee note is very minimal.

Outcome 1: If Alltaxes including income tax is removed:

- Salaried people will bring home more money which will increase purchasing power of the family.
- All commodities including petrol, diesel fast moving consumer goods(FMCG) will be cheaper by 35% to 52%.
- There is no tax evasion and there will not be any black money generation.
- Business sector will be boosted and self-employment will increase.

Outcome 2: When Banking Transaction Tax (BTT) is implemented:

- If BTT is implemented Government can generate more revenue. For example, if BTT is fixed as low as 0.7% to 1%, it could boost banking transaction many fold.
- No separate machinery like income tax department is needed and tax amount can be directly deposited in State/Central/District administration account.
- As transaction tax amount will be very less, public will prefer it instead of paying huge amount against direct/indirect taxes.
- There will be no tax evasion and Government will get huge revenue for development and employment generation.

- For special projects, Government can increase BTT for example, from 1% to 1.2%.

Outcome 3: If 1000 and 500 Rupee currency notes are withdrawn from circulation:
(1) Corruption through cash could be stopped.
(2) Black Money will either be converted to white or will vanish as Billions of 1000/500unused currency will become null and void.
(3) Unaccounted hidden cash that increases the prices of immovable properties and commodities will bring down the prices.
(4) Terrorism supported by cash transaction will be stopped.
(5) Will increase the income to government through stamp duty when the transaction for buying and selling properties is routed through the banking systems.
(6) Circulation of counterfeit notes will be stopped as the printing for less value notes will not be profitable.

Effect of "Arthakranti Proposal" if implemented:

- Prices of all commodities will come down
- Salaried people will get more cash in hand
- There will be an increase in the purchasing power of the society
- Demand will be boosted which will increase production and more employment opportunity
- Surplus revenue to the Government for effective health/ education/ infrastructure/ security/ social works
- Lower Interest rates
- Product quality will improve
- Real estate prices will come down
- Increased funding for research and development

RESEARCH SIGNIFICANCE AND FINDINGS

Drawbacks in Infrastructure:

The traditional infrastructure of e-governance acts as a greater hobble to implement the newly proposed work for tax management and black money curtailment. This is due to[1]

- Application Life Cycle management
- Software licensing and support
- Scalability
- Accountability
- Modifiability
- Physical security

Drawbacks in the method used:

The method which is currently followed suffers from a major drawback of accountability and low monitoring.

Accountability: The current system fails to check the responsibility of the citizens to pay taxes for every product and services they purchase.

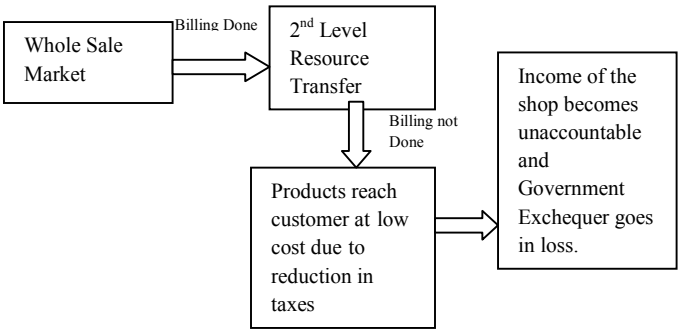


Fig. 3 Current System

Low monitoring: The present system does not track the expenditure of an individual**.** This results in illegal clubbing of money in few hands.

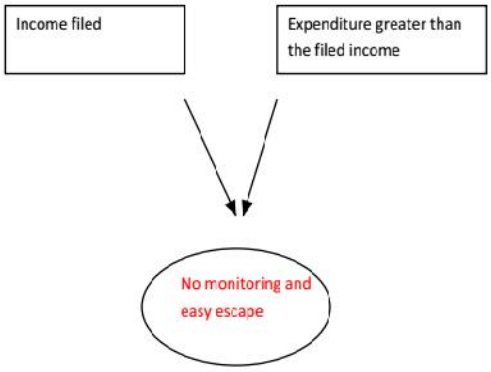


Fig. 4 Major drawbacks in the current system

PROPOSED SYSTEM AND SUGGESTIONS FOR FUTURE RESEARCH:

Cloud as the base:

Cloud offers standard platforms in terms of providing different kinds of systems, middleware and Integration systems. Some of the standard platforms they provide are:

- OS provisioning
- Queuing Service
- Database Services
- Middleware Services
- Workflow Services

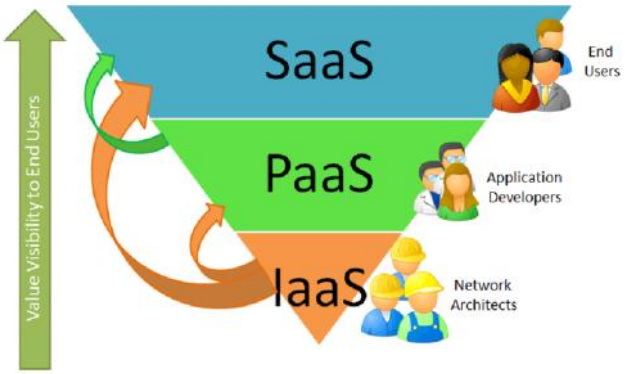


Fig. 5 Cloud services

Government departments requiring resources can get resources instantly as compared to traditional methods where they have to wait till they are purchased and deployed. Applications requiring middleware services can be provided instantly.
Cloud offers applications as a service. For example, a district decided to move to E-governance solution for some application for their citizens need not purchase applications, hardware and software. They can make a request for a particular service from the cloud provider. Applications instances can then be created for their use. Numerous applications can be provided as standard services, where departments can request and manage E-governance applications. Cloud computing supported E-governance can provide efficient management and disaster recovery.

Money Card in Cloud:

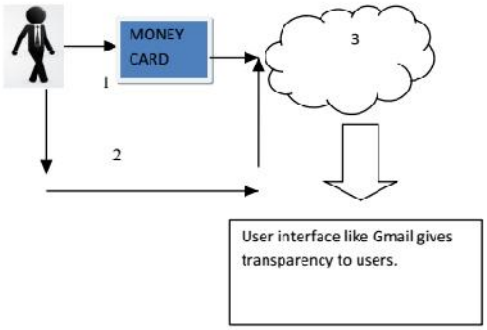


Fig. 6 Schematic representation of money card

1. Any income or purchase will have to be done only through a card based system. Individuals are not allowed to have cash more than 10 % of their income.
2. Since the card cannot be used at all times, cash is used but bills have to be obtained for the purchase will have to be updated in the individual's account.
3. All transactions are accounted in the software loaded in the cloud.
4. Functions of saas:
   - The software in the cloud plays a major role in maintaining individual's income and expenditure.
   - Software is designed based on the government's norms to calculate the taxes automatically, based on the income and expenditure.

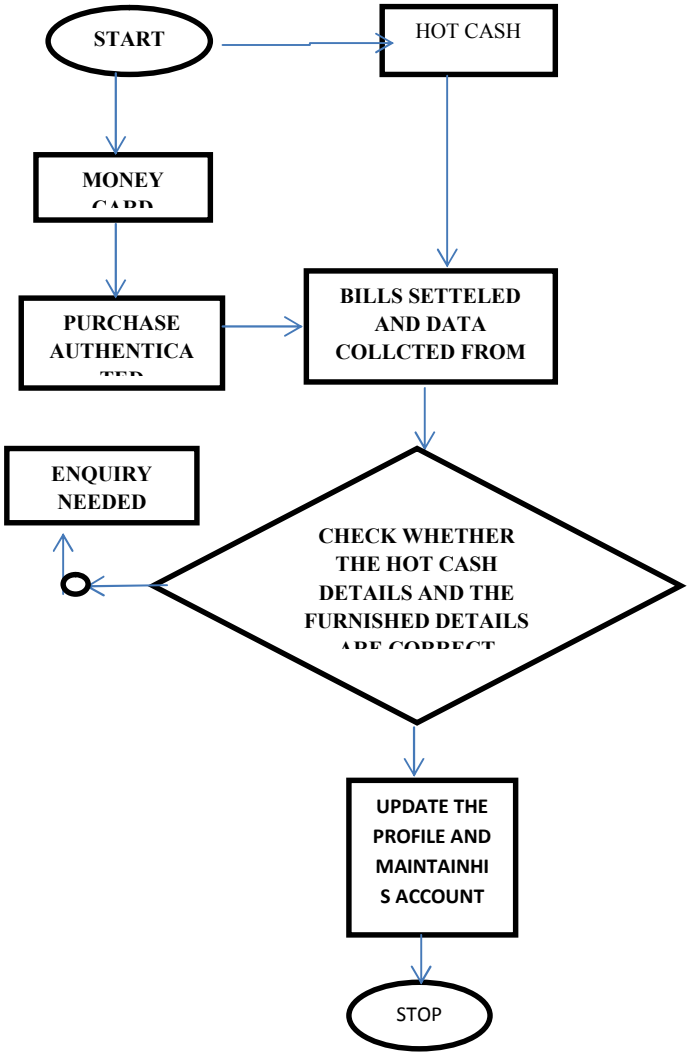


Fig. 7 Flow of proposed system

- For cash transactions bills need to be provided. This can be done through the interface provided or through e-filing.
- The software checks the cash against expenditure.
- The tax detected is automatically reverted to the governments exchequer

***IMPLEMENTATION IN CLOUD COMPUTING**:*

Windows Azure uses a specialized operating system, called Windows Azure, to run its "fabric layer" — a cluster hosted at Microsoft's data centres that manages computing and storage resources of the computers and provisions the resources (or a subset of them) to applications running on top of Windows Azure. Windows Azure has been described as a "cloud layer" on top of a number of Windows Server systems, which use Windows Server 2008 and a customized version of Hyper-V,[1] known as the *Windows Azure Hypervisor*[2] to provide virtualization of services.[1] The platform includes five services - Live Services, SQL Azure (formerly SQL Services), AppFabric (formerly .NET Services), SharePoint Services, and Dynamics CRM Services [3] - which the developers can use to build the applications that will run in the cloud. A client library, in managed code and associated tools are also provided for developing cloud applications in Visual Studio. Scaling and reliability are controlled by the Windows Azure Fabric Controller so the services and environment do not crash if one of the servers crashes within the Microsoft data centre and provides the management of the user's web application like memory resources and load balancing. Windows Azure can currently run .NET Framework applications compiled for the CLR, while supporting the ASP.NET application framework and associated deployment methods to deploy the applications onto the cloud platform. It can also support PHP websites. Two SDKs have been made available for interoperability with Windows Azure: The Java SDK for AppFabric and the Ruby SDK for AppFabric. These enable Java and Ruby developers to integrate with AppFabric Internetservices.Access to Windows Azure libraries for .NET, Java, and Node.js is now available under Apache 2 open source license and hosted onGitHub. A new Windows Azure SDK for Node.js makes Windows Azure an excellent environment for Node applications and a limited preview of an Apache Hadoop-based service for Windows Azure enables Hadoop apps to be deployed in hours instead of days.

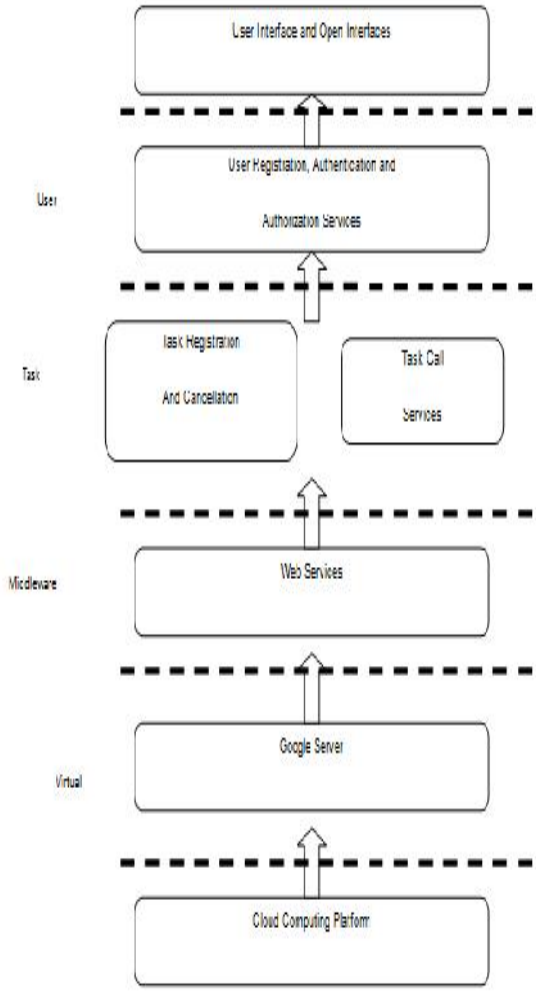


Figure 8 Implementation sample

ADVANTAGES:

This system provides an efficient way of monitoring the taxes and account details. It empowers a corruption free society. This method overcomes all the drawbacks of the present system and provides high fidelity.

- Simplified user interaction
- Provisioning enables policies to lower cost
- Increase system administrator productivity
- Improve service delivery to citizens
- Automate virtual infrastructure for peak performance
- Reduce costs

CONCLUSION:

This study introduces a conceptual model for e-governance. The proposed *THE MONEY CARD SYSTEM* to curtail the flow of unaccounted money can be effected to a great extent. Cloud provides a solid foundation for the introduction of widespread provision of services to various stakeholders. Applications designed using the principles of Service Oriented Architecture deployed in cloud architecture will benefit the government in reducing operating costs and increasing the governance. SOA and cloud architectures when properly applied to developing E-governance applications have the capability to transform the nation into an Information Society.

Service Level Agreement will indicate to measure how well the services are being performed. Cloud helps to enable various e-government services faster and cheaper thereby accelerating the adoption and use of Information Technology for e-services. Cloud architectures allow rapid deployment of turnkey test environments with little or no customization. This work my provide a basis for future research in the associated areas.

REFERENCES:

[1]KeithWard. "More Azure Hypervisor Details" Virtualization Review.Retrieved 16 April 2009.
[2]Alessandro Perilli. "Windows Azure uses a hypervisor but it's not Hyper-V". Virtualization Info.Retrieved 16 April 2009.
[3]Cesar de la Torre (2 November 2008). "Windows Azure".Microsoft.Retrieved 18 November 2008.
[4]ArthaKranti. www.arthakranti.org.

SESSION

E-BUSINESS, E-COMMERCE, ENTERPRISE INFORMATION SYSTEMS, AND SECURITY

Chair(s)

TBA

How to create an E-Advertising Domain Model: the AEADS approach

Alaa A. Qaffas, Alexandra I. Cristea
Department of Computer Science
The University of Warwick
Coventry, CV4 7AL, UK
{aqaffas, acristea}@dcs.warwick.ac.uk

Abstract—**E-advertising is an increasingly profitable industry that continues to grow rapidly year upon year. Correspondingly, the number of people accessing the internet each year continues to rise. However, users respond negatively towards web based advertising campaigns; a prior study suggests that businesses should consider personalized web based advertisements as key to maximising the effectiveness of advertisements on their target customers. To this end, we are developing a toolset to support both authors and customers of adaptive advertising. Whilst personalisation is desirable, the creation of such systems is non-trivial. Thus, in this paper, we introduce a domain model tool, which is the main tool for authoring adaptive advertisements. It can be used by business owners to organise, label and categorise advertisements. Moreover, this tool has been evaluated by companies in the United Kingdom and Saudi Arabia.**

Keywords-E-advertising, E-commerce, Personalisation, Adaptive Advertising, Domain Model, Authoring System.

I. INTRODUCTION

E-advertising experienced a market growth rate of 18% over 2011 and 2012, and as, e.g., the Interactive Advertising Bureau [12] has shown, the industry's income expanded from $7.8bn to $9.26 in this period. It is thus clear that e-advertising is a sector experiencing significant growth [12].

Initially, websites were split between online retailers and those which displayed other content but generated revenue through advertisements. These advertisements were usually simple banners during the late 1990s. Nevertheless, sites have now realised they can make more profits by collecting demographic information about their users through subscriptions. Indeed, the methods used to grab the attention of users have become more aggressive – advertisements now use sounds and even objects that cover the screen. Perhaps unsurprisingly, the frustration that users often feel when confronted with these adverts has led to the development of a range of advert-blocking software [2].

Studies suggest that websites offering personalized content attracted 80% of Internet users to the sites they visited since 2005 [7]. From this, an opportunity has opened up for personalised electronic advertisements. In order to achieve this, it is necessary to monitor the online behaviour of individuals, ensuring that targeted advertisements reach them in the most effective way.

Thus, the main questions this research aims to address are:

1) How can we support the creation of adaptive advertising by website owners?

a) What type of tools do website owners need to be able to efficiently add adaptive adverts in a lightweight manner (as an add-on) to their website?

b) What kind of support do website owners need to be able to use these tools?

This paper responds to the questions above in a constructive manner, by proposing a set of tools for creation and authoring of adaptive advertising. It then focuses on the main tool of this toolset, the domain model tool, which it implements and evaluates with the help of real business owners.

The following sections discuss related research, domain model implementation and evaluation and provides a conclusion and description of future work.

II. RELATED RESEARCH

Closely related fields are that of authoring of adaptive hypermedia. Adaptive hypermedia systems [4] represent an opportunity to increase personalisation, supplying users with reports on matters within specific areas of interest. This technology helps customers by improving the efficiency and accuracy of the delivery of information [4]. When links provided to other websites or content are altered for the individual, in order to create a more tailored experience, this is referred to as adaptation. The different types of adaptation are referred to as 'navigational' and 'presentational' [15]. An authoring system is a computer-based system used to create, e.g., adaptive web content [9]. Most authoring system for adaptive hypermedia use separate tools for creating domain model (DM), goal and constraints model (GM), user model (UM), adaptation model (AM) and presentation model (PM) [11].

Along with the user model, the domain model is considered one of the main parts of adaptive hypermedia. It is used to describe and categorise all the information content and knowledge accessible in the hypermedia. In general, the structure of the domain model in hypermedia systems are either

hierarchical authoring models, or graphical models [5] that represent pieces of knowledge.

In [1] a hypertext document that automatically adapts to the ability level of the reader uses a simplified form of the domain model without any links between concepts. By contrast, in [8] the domain model contains a hierarchy of concepts, along with details of the attributes and relations between these concepts.

[6] introduces an Adaptive Web Content Delivery System then the domain is all web contents on the internet. With respect to domain model in advertising adaptation systems, the domain model must be representing the available advertisements. In addition, these advertisements must be categorised and divided into groups and subgroups.

For interoperability with other systems, such as delivery systems, or other authoring systems, some of the adaptive hypermedia systems have proposed using semantic web languages (mainly XML) for the internal representation of the various authoring tools [10, 16]. For this reason, XML was also used as internal format for the tool presented in this paper.

III. AUTHORING TOOL FOR E-ADVERTISING

A. *The overall authoring of adaptive e-adverts*

1) Domain model that can be used by business owners to organise, label and categorise advertisements, which will be described below.

2) Adaptation model, which will enable businesses to adapt the advertisements they have organised using the domain model tool to their customers' needs. The model, described elsewhere [14], is not further detailed here.

B. *The domain Model*

The domain model is the one of the main two tools of the AEADS authoring toolset. It includes data and advertisements and how they are organised and classified. The domain model includes groups and subgroups and the end of the leaves are the links to the organised advertisements as can be seen in Fig.1. These are the advertisements that will be allocated on the website host.

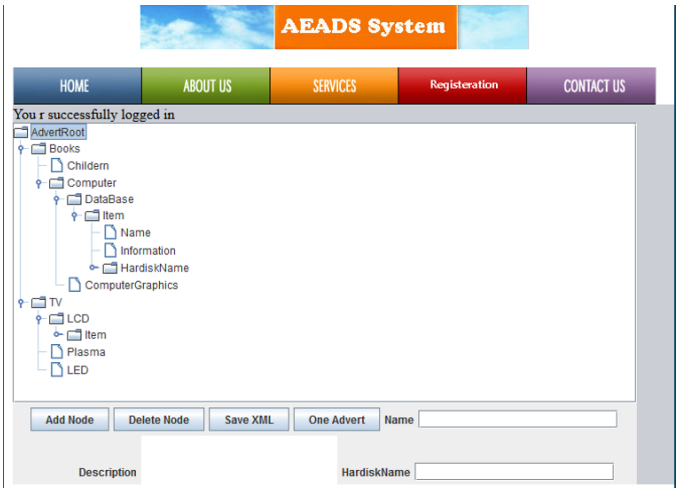


Figure 1. Domain Model Creation

The adverts appear to the business owner as a graphical tree from which to organise and classify ads in simple way. Groups and subgroups can be added and deleted. In addition, each item or advertisement will include a property name, a description and a hard disk name, which will indicate the advertisement's name, any information about the advertisement and the advertisement's name on the hard disk respectively. For internal storage and exportability, an XML file is created for each business owner who wants to organise the advertisements through our system upon registration, as shown in the example in Fig. 2.

```
<?xml version="1.0" encoding="UTF-8"?>
- <AdvertRoot>
  - <Tv>
    - <LCD>
      - <LED>
        - <Item>
            <Name>ad1</Name>
            <Information>TV ad1</Information>
            <HardiskName>TV1</HardiskName>
          </Item>
        - <Item>
            <Name>advert2</Name>
            <Information>TV ad2</Information>
            <HardiskName>TV2</HardiskName>
          </Item>
        </LED>
      </LCD>
    - <Plasma>
      - <Item>
          <Name>ad3</Name>
          <Information>TV ad3</Information>
          <HardiskName>TV3</HardiskName>
        </Item>
      - <Item>
          <Name>ad4</Name>
          <Information>TV ad4</Information>
          <HardiskName>TV4</HardiskName>
        </Item>
      </Plasma>
    </Tv>
</AdvertRoot>
```

Figure 2. XML Sample

IV. CASE STUDY

A. *Hypotheses*

The following hypotheses have been written to evaluate the domain model tool:

H1: The tool is important for our business.

H2: The GUI of the tool is attractive.

H3: The tool makes our work easier

H4: The tool is sufficient for creating and organising all of our advertisements

H5: The tool saves us time

H6: The tool can be used by any website to create and arrange advertisement domains

H7: New staff can understand and use this tool with minimal training

H8: The domain model home is useful and easy to use

H9: Registration is useful and easy to use

H10: Login is useful and easy to use

H11: The creation functions are useful and easy to use

These hypotheses were tested by surveying selected businesses and analysing their responses, as described below.

B. Case Study Setup

The domain model has been tested and a questionnaire has been created, based on implementation and the hypotheses, for business owners to evaluate its ease of use and utility.

The domain model tool was presented to twelve business owners for evaluation, selected especially from a variety of company types. The procedure was as follows.

Firstly, they were informed about the system as a whole, as well as to the idea of adaptive advertising in general. At the end of this presentation, we asked each business owner to use the tool. Then, we asked them to fill in a questionnaire, which included four parts. The first part concerned demographic information. The second part consisted of general questions about tool usability and to what extent the business owners agreed that this tool is important and it makes their work easier. Likert scale [13] questions were used in the third part to get business owners' feedback on tool features and functions, which can be seen in Table 1, and which are further used in Figures 6 and 7. The Likert scale used in this study took the format of a five point Likert scale. Respondents were asked to choose from five answers evaluating the tool's usefulness and ease of use with 1 being not useful at all or very hard to use, and 5 being very useful or very easy to use, respectively. The last part of this questionnaire consisted of open question to obtain any further comments the owners may have had.

TABLE I. KEY FEATURES AND FUNCTIONS

A	Domain Model - Home	K	Login Process
B	Registration	L	Reset Password
C	Login	M	Adding Category - Subcategory
D	Creation Functions	N	Removing Category - Subcategory
E	Logout	O	Adding Advertisement inside subcategory
F	Registration Process	P	Adding Advertisements Name
G	Sufficient Data	Q	Adding Advertisements Description
H	Reset Information	R	Adding Advertisements file name
I	Submit Information	S	Saving the Tree into XML
J	Creating Account	T	Load the XML file(Domain Model) as tree

C. Results

Responses were obtained from businesses in the communication, construction, consulting, media, online education, trading, training and transportation industries (see Table 2). A total of 42% of businesses were classified as small, 33% as large and 25% as medium-sized enterprises (see Fig.3). A total of 58% of the businesses were located in Saudi Arabia and the remaining 42% in the United Kingdom (see Fig.4).

TABLE II. TYPE OF BUSINESS

Type of businesses	Number
Communication	3
Constructing	1
Consulting	2
Media	1
Online Education	1
Trading	2
Training	1
Transportation	1
Total	12

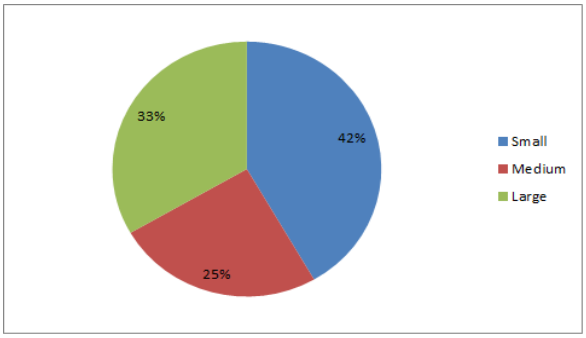


Figure 3. Size of Business

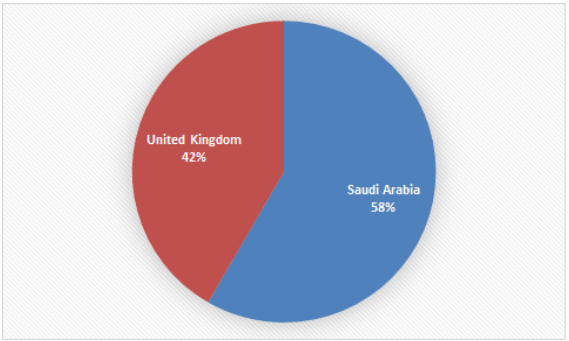


Figure 4. Country

The general questions section included seven questions, alternating between positive and negative tone, to eliminate any bias that could be introduced by the questionnaire [3]. As can be seen in Fig.5, all businesses agreed that the tool is important for their business and makes their work easier in terms of organising their advertisements. In addition, they strongly agreed that the domain model tool saved them time and new staff could understand and use it with minimal training.

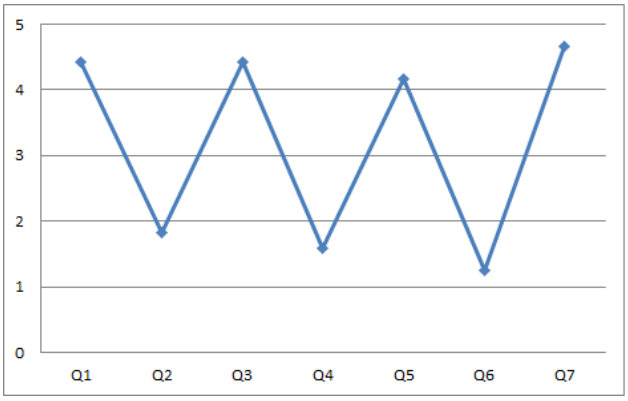


Figure 5. General Questions

As shown in Fig.6, businesses responded to Likert scale usefulness questions for each of the domain model tool's specific features and functions. Following the analysis and tabulation of the data, it can be seen that each of the domain model tool's twenty elements rated very highly. Overall, the results were shown to be between 4 and 5, meaning that the tool is useful, with 4 indicating that a feature or function was useful and 5 indicating that a feature or function was very useful. Scoring highest in terms of usefulness were the registration process and creating an account with the login process following close behind. The lowest scoring features of the domain tool were the reset information and domain model home, although both elements still scored above 4. The possible reason for this is that the implementation of GUI was not complete. So, they felt that these functions are not useful enough. The mean values were greater than 3, all of them were between 4.08 and 4.91, which means the tool is easy to use. The standard deviation values were between .29 and .79.

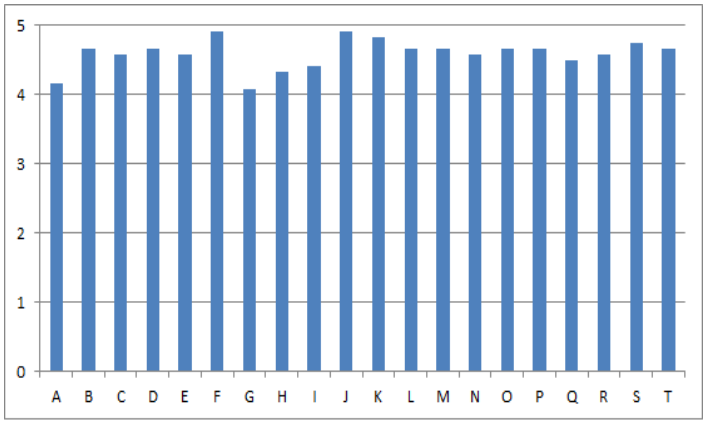


Figure 6. Usefulness (Ox axis detailed in Table 1)

The Likert scale ease of use items were responded to positively by businesses for all of the features and functions of the domain model tool, revealing that it is easy to use, as seen in Fig.7. As with the usefulness items, the ease of use items addressed each of the twenty features and functions of the domain model tool. Again, each feature and function received a score of between 4 and 5. Following the analysis of the data collected in the questionnaire it was shown that registration, login and the registration process all scored perfect 5s. Loading the XML file (domain model) as a file scored lowest but still scored above 4 as it takes a few more time to be loaded, indicating that it was easy to use. These figures, and those recorded above, are very encouraging regarding the future of the domain model tool. Moreover, The mean values were between 4.25 and 5.00 and the standard deviation values were between .00 and .75. The tool is easy to use as the mean values were all greater than 3.

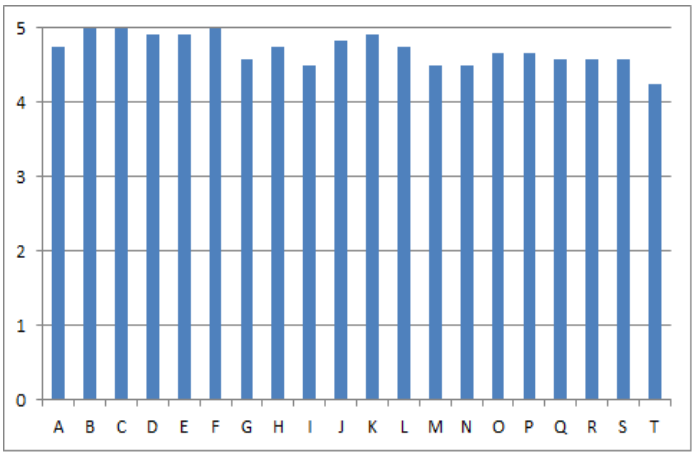


Figure 7. Ease of Use (Ox axis detailed in Table 1)

From the open questions, we note that business owners want to improve the interface of the tool stating that we should: "Improve the GUI, add the actual ads and a preview / Thumbnail" and "It could be more flexible such as changing colours because the colours are very important with advertisements". In addition, owners want the opportunity to write instructions on the first page. Moreover, they were worried about the problems that they may face during the classification process. For example, one wrote: "Is there a support category within this web tool to help the users if they face a problem?" and "If I make a mistake during the classification process and I want to add another subgroup I have to delete items and add this subgroup and re-add items which is a waste of time". The responses to the open questions are very valuable for the furtherance of the domain model. The open questions allowed business owners to make comments outside the more structured elements of the questionnaire and to highlight issues that the researchers may not have been immediately aware of. This can then stimulate new ideas for the tool and will thus enable the researchers to modify the domain model so that it can reach its full potential and offer users the best possible solution for the organisation and categorisation of advertisements. For example, the suggestion to incorporate a support facility will improve user experience.

V. Conclusion and Future Work

We believe that an adaptive system will help businesses to increase their revenue by enabling them to send the appropriate advertisements to the appropriate customers at the right time. The first tool of the AEADS has been implemented to allow businesses to organise their advertisements inside groups and subgroups and to attach any necessary information for these advertisements, which makes their work easier and saves them

time. The features and functions of this tool have been tested. Furthermore, the tool has been evaluated by business owners and they have a positive attitude towards it with all of the features and functions of the domain model tool scoring between 4 and 5 on a Likert scale in terms of their usefulness and ease of use (with 1 being not useful all and not easy to use and 5 being very useful and easy to use). However, the open questions section of the questionnaire indicated some of the concerns that business owners had about and suggestions for the domain model tool which need to be addressed. The next tool to be included in this system will be an adaptation strategy tool which will enable businesses to adapt the advertisements they have organised using the domain model tool to their customers' needs. This will allow businesses to personalise their advertisements and target their promotional materials more effectively and thus increase their profitability.

REFERENCES

[1] C. Boyle, and A. O. Encarnacion. "MetaDoc: an adaptive hypertext reading system." In *Adaptive Hypertext and Hypermedia*, 71-89: Springer, 1998.

[2] M. Brain. "How Web Advertising Works." (2002). http://computer.howstuffworks.com/web-advertising.htm [accessed 10 March 2014].

[3] J. Brooke. "SUS-A quick and dirty usability scale." *Usability evaluation in industry* 189, (1996): 194.

[4] P. Brusilovsky. "Adaptive hypermedia: An attempt to analyze and generalize." In *Multimedia, Hypermedia, and Virtual Reality Models, Systems, and Applications*, edited by P. Brusilovsky, P. Kommers and N. Streitz, 1077, 288-304: Springer Berlin Heidelberg, 1996.

[5] P. Brusilovsky. "Methods and Techniques of Adaptive Hypermedia." *Adaptive Hypertext and Hypermedia*, (1998): 1.

[6] J. Chen, Y. Yang, and H. Zhang. "An Adaptive Web Content Delivery System." In *Adaptive Hypermedia and Adaptive Web-Based Systems*, edited by P. Brusilovsky, O. Stock and C. Strapparava, 1892, 284-288: Springer Berlin Heidelberg, 2000.

[7] ChoiceStream. "ChoiceStream Personalization Survey: Consumer Trends and Perceptions." (2005). http://www.choicestream.com/pdf/ChoiceStream_PersonalizationSurveyResults2005.pdf [accessed 10 March 2014].

[8] A. Cristea, D. Floes, N. Stash, and P. De Bra. "MOT meets AHA!", (2003).

[9] A. Cristea, and C. Stewart. "Authoring of Adaptive Hypermedia." In *Advances in Web-Based Education: Personalized Learning Environments*, edited by G. D. M. a. S. Y. Chen, 8, 225-252: Information Science Publishing (IDEA group), 2006.

[10] A. I. Cristea. "What can the semantic web do for adaptive educational hypermedia?" *Educational Technology and Society* 7, no. 4 (2004): 40-58.

[11] A. I. Cristea, and A. de Mooij, "LAOS: Layered WWW AHS authoring model and their corresponding algebraic operators," in WWW03 (The Twelfth International World Wide Web Conference), Alternate Track on Education, Budapest, Hungary, 2003.

[12] InternetAdvertisingBureau. "Internet Advertising Revenues Hit Historic High in Q3 2012 at Nearly $9.3 Billion, Rising 18% Over the Same Period Last Year, According to IAB." (2012). http://www.iab.net/about_the_iab/recent_press_releases/%20press_release_archive/press_release/pr-121912 [accessed 10 March 2014].

[13] J. McIver, and E. G. Carmines. *Unidimensional scaling*. Vol. 24: Sage, 1981.

[14] A. Qaffas, and A. Cristea, "How to Create an E-Advertising Adaptation Strategy: the AEADS Approach," in The 15th International Conference on Electronic Commerce and Web Technologies(EC-Web'14), Munich, Germany, 2014 (Accepted).

[15] A. Smith. "MLTutor: A Web-based educational adaptive hypertext system." Ph. D. dissertation, School of Computing Science at Middlesex University, 1999.

[16] H. Wu. "A reference architecture for Adaptive Hypermedia Applications." Technische Universiteit Eindhoven, 2002.

A Three Tier Enterprise Information System

S. Duggineni[1], D. Yoon[1]
[1]CIS Department, University of Michigan – Dearborn, Dearborn, MI, USA

Abstract – *With rapid development of distributed systems, the characteristics of enterprise information systems are changing along with them. The classical enterprise information system is based on the master-slave model, the two-tier models have emerged based on the client-server systems, and multi-tier models have appeared along with internet-based systems. In this paper, we will present a three-tier enterprise information system.*

Keywords: Smart Grid, OLTP, OLAP, SQL

1 Introduction

Initially IT systems are used to handle and query transactional data for day to day operations which are called online transactional processing (OLTP) [1] systems. OLTP serves only for transactional execution. To get a high-level summarized data, transactional data needs to be summarized across multiple dimensions and resulting key figures. For example questions like sales across branches, trends across months, regions and to analyze how effective a sales promotion worked to improve sales operations. These kinds of requirements are answered using Online Analytical Processing Systems (OLAP) [2] systems. Enterprise applications for commerce and government are becoming more and more critical and complex day by day and demands both *transactional* and *analytical* capabilities in a single platform. Organizations are trying to simplify Information Technology Landscape, Webifying and moving towards Mobile based applications and getting integrated with social applications. "Enterprise Application Architecture" as a strategy is very important to reduce application maintenance cost and to move business applications to changing customer needs. **Application Architecture strategy needs to ensure the applications should align with the growth strategy of the organization**. For example, if an organization wants to provide mobile based front end applications, it should be able to provide it without doing significant changes in core application framework.

The more complex a business becomes, the more number of databases are designed with a target to handle faster insertions/deletion/updations and analytical and planning applications are taken out from transactional systems to achieve flexibility, performance and to handle data from multiple sources. Using normalization techniques, OLTP systems are designed such a way that less amount of data is queried or inserted/modified to achieve fast response time where as OLAP is designed with star schema as base line in which data is stored in a summarized format to answer multi dimension queries. Even though there is a great deal of research and success is achieved with reducing costs of main memory, increasded multicore CPU's (For Parallel Querying) [7] to combine both OLAP and OLTP in single database environment, at the time of writing this paper still it is in its begining and very costly to implement. Not only high cost of this type of environment, limitations of parallel updates to same database at a single instance of time, database deadlocks [8] during execution of parallel transactions keeps OLTP and OLAP in seperate systems. Even though it is possible to query data from multiple databases using ODBC [9], OLE DB [10], JDBC [11] etc, different SQL engines implement SQL slightly differently and record results needs to be handled differently based on underlying database provider. Also application developer needs to know which data is available in OLTP and which one is available in OLAP. In addition to this, application developer should remember the SQL syntax for different SQL engines. This adds significant overhead in application development and also portability issues when underlying database is changed. There is no intelligent mechanism available for when to query data from OLTP and when to query from OLAP. Many OLAP database providers offer standard DB level operations to work on data in DB server itself. This feature needs to be used efficiently without worrying about underlying syntax.

Not only different databases adding to the complexity, frequent changes of database design also causing the application layer to be re-designed frequently. In most of the enterprise applications data is stored in third normal form (3NF) [12] but sometimes due to "data selection vs insertion analysis" over time, results to re-organizing of existing logical models (Large Joins vs Fast Insertions). This will lead to normalizing or de-normalizing of physical data models and application needs to adopt it automatically without changing much of its operating behavior. With these problems, an enterprise architectural approach is needed to combine OLTP and OLAP with existing infrastuctures, accomodating database changes dynamically, intelligent query mechanism between OLTP and OLAP, application querying without worrying underlying database, using business context information for querying.

2 Multi Tier Architecture and Its Importance

To achieve scalability, robustness and flexibility, IT applications moved from single tier applications to two tier and three tier-applications. Back in the days of huge main frame systems, all processing is executed on a single computer node, which is an example of single tier system. In single tier system, all computing resources are attached in a

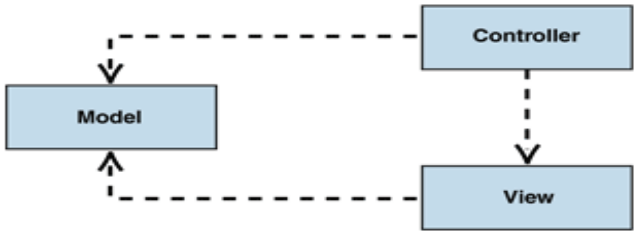
single instance and user interface is achieved like telnet or dumb terminals. Even though it is a very simple architecture and very efficient in performance (no need to transfer the data between multiple layers), it is very expensive and also the flexibility and scalability is limited. From here, two tier architecture is used with evolution of browsers where rendering of data is possible. Dual tier architecture is used with personal computers and connecting to servers when needed. This is often called like a client-server model. The main advantage of dual tier computer is less expensive than main frame and offline applications can be executed (Ex: Store data on personal computer and upload to server later). The main disadvantage of this is, both database and applications are handled in a single node and scalability of database operations is limited. This lead enterprise systems to use the three tier architecture [13]. Using three tier architecture offers the following advantages.

1. Application server can be deployed on many machines.
2. It is no longer needed to maintain separate database connections for every user.
3. In addition to the above, achieving data integrity is easy with three-tier as application server needs to validate that only valid data is allowed to enter into database.
4. Upgrades are easy and less down time for mission critical applications.

3 MVC Architecture Pattern

Many enterprise applications and specially e-commerce applications are built using three tier architecture and achived scalability and other advantages as discussed above. But clients in real world commercial applications expects that data needs to be displayed in different formats. Development of different screens and adjusting application layer for each front end requirements is a tedious task. For example, a trading system needs to be accessed from mobile, desktop rich clients, browsers, PDA's. Same data needs to be displayed in a dashboard or list format. With the advance of programming languages rapid GUI building capabilities, building user interfaces became easy. Tradiotionally developers used to tie UI generation/handling events and application processing logic. This helps to reduce the amount of coding (Ex: Same program handles both UI actions and Processing logic) and gave better performance but the flexibility of using and showcasing data independent of UI is lost. Applications built this way cannot meet fast changing ways of consuming data from different devices in different formats. Ex: Users bank account login from various devices. To solve this problem MVC Architecture pattern needs to be used.

In MVC pattern [14], the user input, the modeling of the external world, and the visual feedback to the user are explicitly separated and handled by three types of object, each specialized for its task. The view manages the graphical and/or textual output to the portion of the display that is allocated to its application. The controller interprets the mouse and keyboard inputs from the user, commanding the model and/or the view to change as appropriate. Finally, the model manages the behavior and data of the application domain, responds to requests for information about its state (usually from the view), and responds to instructions to change state (usually from the controller).



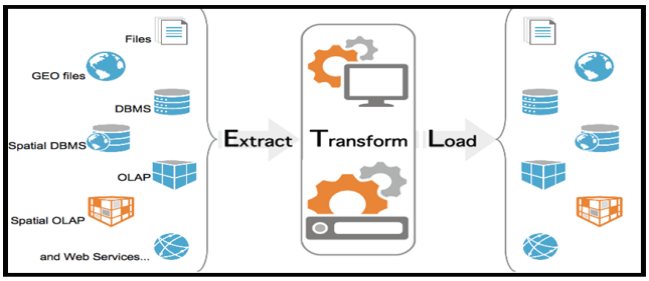
4 OLTP design and its importance in efficient transactional processing

The main goal of any transactional system is to collect/retrieve, process and store data as efficiently as possible. In order to achive fast response times, the amount of data transferred between application layers and database layers should be minimal. This lead to normalizing of database tables. Normalization [15] of database is very important to avoid redundant storage of data. If database tables are not normalized, transactional systems will face performance issues during high degree of insert and update statements. If tables are normalized, more number of joins are needed to retrieve the required data and often this leads to de-normalizing decissions based on the way data is accessed.
Based on application requirements, good database designing and normalization principles need to be followed to achive fast database response times. E-R modelling needs to be used to represent a logical model of database and this should acts as a base line to construct physcial model.

4.1 Normalization of database

The aim of transactional database design is to minimize the data storage for master and transactional data and eliminates unnecessary data redundancy. With data normalization any tabular -columns dependencies are detected and the table is restructured into multiple tables (two tables) which eliminate any column dependency. Incase the data dependency is still exhibited, the process is repeated till such dependency is eliminated. The third normalization form produces well designed database which provides a higher degree of independency.

Ex: {Cid, Cname, Hno, Street, City, State, Zip }==>{Cid, Cname, Hno, Zip } & {Zip, City, State }

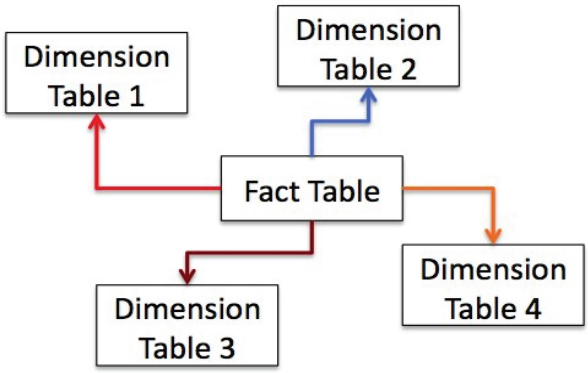


5 OLAP design approach, importance and design principles

To stay competetive in business, it is very critical to have a analytical view of transactional data i.e generated over time. Only with right analytical inputs, management teams can gain important business insights to achieve business goals. Analytical applications are more concentrated towards summarized data rather fine granular data. Many times, analytical applications needs to get data in a uniform manner from different sources or from different formats. For example, transactional data might be in different currencies but analysts want to work with uniform currency. To get data from different datasources and transform into a uniform manner, ETL (Extraction-Transformation-Loading) [16] tools needs to be used.

For analytical applications, Query throughput and response times are more important. In analytical applications users always want to analyze key figures based on dimensions. For example, questions like sales across branches, sales across products, sales for product and branchwise.

To answer these kind of questions, OLAP applications are designed using star schema approach [17]. Main purpose of star schema is to identify dimensions on which we want to analyze the data and the numbers what we need to analyze which are called key figures. Any OLAP design starts with identifying dimensions, facts and to what level of aggregation the data is getting stored in data warehouse. Deciding the required level of aggregation is a key part of the speed of cube based reporting. The reason why a cube can be very fast, when for example selecting data for an entire year is because it has already calculated the answer. Whereas a typical relational database would potentially sum millions of day level records on the fly to get an annual total, Analysis Services cubes calculate these aggregations during the cube build and hence a well designed cube can return the answer quickly. Deciding what level of aggregation based on what level of granularity is needed in reporting.



6 Why we need answers from both OLTP and OLAP in real time

Till date, most of the enterprise applications are built in such a way that separate application platforms are built for OLTP based applications and separate for OLAP based applications. OLTP is used in day to day business operations like creating sales order, credit or debit transactions like these. OLAP is used for management information systems which help in decision making based on summarized data from multiple data sources and is usually used by the management teams. Even though OLTP and OLAP is designed for different purposes, with increasing competetion and customer expectations transactional users needs to have applications with both analytical and transactional capabilities. If transactional users has real time business intelligence, then business process can be optimized dynamically. **Having real time business intelligence is going to be one of the main component of future business applications**. Main applications of real time analytics or side by side OLTP and OLAP applications include in areas like Customer Centric Personalized Promotion Management, Targeted Live Chat with customer where chance of making a deal is 0-100% based on personalized price, mobile based ads (Only related ads: Best example Google's ad sense, Google Trends), smart electric grid etc.

These kind of smart business applications are already showing good results in retail market.

Due to primary constraints of OLTP and OLAP, seperate databases are used to store and process transactional and analytical data and till we get an affordable common database for OLTP and OLAP, existing IT applications will store OLTP and OLAP data seprately but they no need to have seperate application layer to showcase and process the data.

Earlier separate application layers are built for Transactional and Analytical Applications. The main reason for this is, lack of ready to use graphical display capabilities in old programming languages,not many standard pre-delivered analytical libraries, in-built predictive analytic functions at database level were not available and the efforts involved in customizing the libraries. This made organizations to build separate application platforms for OLTP and OLAP Applications. With the advance of programming languages,

browser capabilities, in built predictive analytics embedded in the database, it is no longer needed to have separate application platforms handle OLTP and OLAP. Having separate application platforms are adding considerable overhead cost and performance issues to organizations. **Future e-commerce applications needs to be designed to have access to both analytical and transactional features in a seamless way without worrying about underlying database**. In future, any database changes happen or if common database approach is used, it should be easy to migrate without changing the application platform.

7 Importance of database independent application layer

IT systems for large organizations often involve different databases and often implemented on differenet set of architectures. Even though having different databases causes high maintenance cost due to duplication of efforts, these are very common. For example if two large organizations merge, they would like to use two seperate database entities but want to use common application platform to reduce application maintenance cost and to achieve uniformness across organizational entities. If applications are designed or developed for a specific database instance, same application might not be used by other organizational entities that use different database. Application Layers interoperability is a very important features that is needed to solve these kind of problems. Different database SQL servers implement different parts of standard SQL. In addition to this, database vendors are constantly adding new inbuilt analytical functions with each database upgrade or release. All these features should be easy to consume by the application without woryying about underlying SQL syntax.

Achieving database portability is very important for any application to survive a long period. Having a database abstraction layer relieves application developer to worry about data retrieval logic and also the application can easily accommodate any database changes. In addition to this, system should intelligently route the constructed queries to either OLTP or OLAP based on available database context.

8 Importance of Database Context and Advantages of Storing it in XML format

For efficient retrieval and process of data, application developer should have knowledge about complete database structure. In many large database structures, it is often a difficult task for application developers to know about all underlying tables involved, fields available, ready to use inbuilt analytical functions that can be used. To solve this problem, complete database context information can be stored in XML [18] format. Application developers refer only database context information and calls database abstraction layer API's that constructs dynamic SQL queries based on business context input, output and analytical operations. Many of times, if underlying tables or fields change, it should be easy to accomodate the changes in application layer. Ex: If normalization or de-normalization happens, application layer should not have any effect.

Main advantages of storing database context in XML is,

i) It is easy to read and understand.

ii) All most all available languages provide application programming interfaces (API's) to process

XML data and due to self-description features using meta tags.

9 Conclusion

Real time analytics in transactional processing systems is a new trend in IT business applications. It is already proving a good success in the market. This paper discusses the importance of multi-tier architecture to achive scalability, OLTP features and why efficient OLTP design is important for fast transactional processing and why OLAP is designed seperately from OLTP to achieve multi dimensional analysis. Importance of real time analytics for new dimension applications and how to integrae OLTP and OLAP in a side by side scenario using the advantage of new dimenstion programming languages capabilities, why database independent application layer is needed for an application to survive a long period, what are the main architecture principles that needs to be followed to build these kind of applications with existing IT infrastructure without additional cost.

10 References

[1] OLTP :http://en.wikipedia.org/wiki/Online_transaction_processing

[2] OLAP: http://en.wikipedia.org/wiki/OLAP_cube

[3] 64 bit architecture: http://en.wikipedia.org/wiki/64-bit_computing

[4] H Plattner A Common Database Approach for OLTP and OLAP using an in-memory column database http://dl.acm.org/citation.cfm?id=1559846

[5] SQL: http://www.w3schools.com/sql/sql_quickref.asp

[6] The Smart Grid:An Introduction| Department of Energy http://energy.gov/oe/downloads/smart-grid-introduction-0

[7] Multicore Processors- A Necessity by Bryan Schauer http://www.csa.com/discoveryguides/multicore/review.pdf

[8] M Grechanik, B.M. Mainul Hossain, Ugo Buy, Wang: Preventing Database Deadlocks in Applications http://www.cs.uic.edu/~drmark/index_htm_files/REDACT_FSE13.pdf

[9] An Introduction to ODBChttp://msdn.microsoft.com/en-us/library/ms715408(v=vs.85).aspx

[10] OLE DB Wiki Referenece http://en.wikipedia.org/wiki/OLE_DB

[11] JDBC Developer's Guide and Reference http://docs.oracle.com/cd/B19306_01/java.102/b14355.pdf

[12] A Simple Guide to Normal Forms in Relational Database Thoery by William Kent-1982 http://www.bkent.net/Doc/simple5.htm

[13] Using a Three-Tier Architecture Model http://msdn.microsoft.com/en-us/library/windows/desktop/ms685068(v=vs.85).aspx

[14] Using MVC in Small Talk < http://st-www.cs.illinois.edu/users/smarch/st-docs/mvc.html>

[15] Codd, E.F. (June 1970). "A Relational Model of Data for Large Shared Data Banks". Communications of the ACM 13 (6): 377–387. http://dl.acm.org/citation.cfm?doid=362384.362685

[16] Conceputla Modelling for ETL processses <http://dl.acm.org/citation.cfm?id=583890.583893 >

[17] Star Schema The Complete Reference by Christopher Adamson ISBN 978-0071744324

[18] XML Essentials http://www.w3.org/standards/xml/core

The new exploration of network marketing based on “S.CN”

Wenrong Jiang, Jiajun Tang
School of Computer and Information, Shanghai Second Polytechnic University,
Pudong New District, Shanghai, China
wrjiang@sspu.edu.cn, 307572027@qq.com

Abstract - *Nowadays, Electronic commerce is very popular, and network marketing which is closely associated with it has been increasingly important. The purpose of this paper is to analyze the advantage and the future development of the network marketing. This article will introduce a successful outsourcing case and analysis of the importance of network marketing. Through the analysis , it means to lead to some better suggestions and methods, and expound the condition and the basis of independent development. Through the case column explains the practical significance of the network marketing, as well as the preparation, methods of network marketing to deployment and implementation and later maintenance and so on. Through the successful case of "S.CN" , I will expounds some widely used and relatively suitable strategy of network marketing. Analyzing the case of "S.CN", I will introduce some methods of network marketing which will be used in the near future, and make some feasibility analysis and detailed implementation investigating. According to the analysis of the situation, a better platform which will be established will make network marketing get better development and utilization. In addition this paper will give some advice and attempts about network marketing and Search Engine marketing,also analyze how to make e-commerce better and to be more linked to network marketing, so that the network marketing activity in the future performs more active and more effective.*

Keywords: network marketing; S.cn; Search Engine Optimization; new exploration.

1 Introduction

On November 11, 2013, an e-commerce storm—“Tao Bao 11.11” is released. This is a very successful case for network marketing, the major e-commerce companies want to share the benefit in that period of time. It may be said to that if your e-commerce activities want to be success, you should match with extremely mature network marketing to it.

Although, informatization has developed in China for many years and the network technology continues to improve, the cognitive deficiency of network marketing makes China Internet market extremely chaotic. With the demand of e-commerce increasing, the network marketing which has great relation with e-commerce is gradually coming into the mind of people. It could say that what is driving the market and makes contributions to build the market order is the success of the network marketing. In China, the number of company which is involved in the network marketing is not a few, also most of them have their own website and they want to make better in e-commerce and network marketing, but the efforts is often insufficient, so that they have to rely on network marketing outsourcing in order to obtain the highest profit return. However to synchronize the resource and standardize the market becomes more and more important.

This article will introduce an e-commerce website—S.CN, and analyze its network marketing activities. With the help of this case, I will put some opinions to the development and innovation of network marketing.

2 Analysis of the present situation of the network marketing

Today's society has entered the network economy era which is based on the Internet. The network technology has comprehensive influence on the contemporary society and brings a huge change to the marketing management for enterprise. Network marketing is a new marketing method which adapts to the development of network technology and the change of information era, it has changed the traditional marketing concept, marketing strategy and marketing methods and it will replace the traditional marketing into the mainstream for marketing means in the 21st century. Network marketing is broad, real-time, economy and interactivity which is traditional lacks for the traditional marketing.

Though network marketing has many advantages, some of the disadvantage of network marketing is also unavoidable. At present, the quantity of the Internet computer and the total number of users are giant, but Internet popularity is obviously insufficient, and the application of information technology in our country enterprise appears extremely not optimistic. Compared with foreign advanced concept, technology, and high popularity, we can find that the overall level of China's network marketing is still in a lower level, the

obstacles and problems which restrict the development of our enterprise network marketing mainly divided into:

2.1 Obstacle for the development of network marketing: knowledge barriers and management barrier

Knowledge barriers is the fundamental obstacle of the development of network marketing in China, companies tend to view of the network marketing is divided into: cognitive deficiency and excessive reliance, two of them is a one-sided view, it is that who cannot be fully understanding of network marketing will bring great harm to the actual operation.

Management barriers mainly refer to loopholes of some enterprises management system that makes management of network marketing extremely chaotic. This way of management is a kind of passive management, and the system generated by the enterprise itself often cannot be implemented in the network marketing activities. That situation weakened the real benefits of the network marketing at one certain extent. Management barriers also are reflected in our unsystematic not enough enterprise management.

2.2 Question for the development of network marketing: weak Internet infrastructure and network marketing talents' lack

Weak infrastructure mainly exists in regional differentiation, although the communication network construction is developed in our country in recent years, due to our country development imbalance and higher communication needs of network marketing to highlight the differences in the network marketing development, this also causes the network marketing activity not being fully development.

The development of network marketing needs all aspects of the talents, especially with the new concept of information and knowledge of information structure. In today's society talent is the lack of resources, thus talents with both technology and marketing become more valuable.

To sum up, though the network marketing is booming in China, many problems hindering the development and the revenue are still existed. And how the network marketing can be in an important position in the electronic commerce activities is relying on its wide range of advantages, it is extremely important how to foster strengths and circumvent weaknesses.

3 Analysis of S.CN 's network marketing

If you cannot find the right of communication strategy in the means of communication revolution caused by information technology, your market will be further squeeze and enterprises will be marginalized. Thus this chapter will introduce a network marketing success case by the introduction of this case, the problem of enterprise and solution, by the case to see the network marketing to bring real benefits and the road to success.

3.1 Introduction of enterprise

S.CN is born on the global recession; it means that opening the designer shoes' warehouse, so the name is: "S.CN". The original LOGO of S.CN is composted by the English word "SKOMART", it is easy to be confused, because the pronunciation is similar to SCHOOL/SCHOOL. Consequently, in July 2010, S.CN opens one of the shortest domain names in the world, domain name owns only single letter S, at the same time, the company name becomes S.CN network technology co. LTD., the LOGO also changed accordingly upgrade (as shown in Figure 1).

Fig. 1. S.CN LOGO and domain name

S.CN reminds every employee that S.CN is a careful and meticulous retailers. In the concept of "value to the guest for honor" of website operation, the customer benefit is always existed in the first place.

In 2008, the founder and CEO of S.CN Songmao XU made S.CN enter the domestic market with the mission of "afford young people to buy authentic designer shoes" and the form of "integrity, pragmatic, enterprising", the total investment of 60 million Yuan was used to build shoes online retailer brand, S.CN has been the leading of authentic designer shoes discount retail sites in only two years.

3.2 Analysis of network marketing

S.CN is an e-commerce site, the success of its network marketing is search engine optimization (hereafter are expressed in SEO), even means a quadratic optimization. Here is to introduce S.CN's success: for network marketing:

- The setting of keyword

Title: S.CN, bargain designer shoes and clothing website, buy shoes in S.CN

Keyword: S.CN, shoes, buys shoes

Description: S.CN is a professional sports shoes and clothing website for NIKE, Adidas, Lining, Anta, PUMA, New balance, Converse, Kappa, and Timberland and so on.

It is a crucial technique for layout that set up good distribution products, product brand, and website brand on the title.

- Brand Keyword

Home page keywords focus on the brand—S.CN, each big brand shoes, brand promotion.

The extremely complex and confusing domain name appears uncoordinated, sitting key words is one of the first consider action of network marketing, and a concise domain name will be easier to make SEO activities.

- Product Keyword

In the baidu searching quality goods can row of the first page, which means this is optimized word which S.CN makes great effort. On the other hand, the rank of searching other brand shoes is not ideal.

In addition, S.CN 's optimization for product model keywords, product classification keywords, product quality description keywords shopping keywords preferential information keywords and so on is highly successful.

Keyword optimization is a part of the network popularization, and also, the website promotion is a kind of network marketing. SEO which is the basis of network marketing is a means of ranking, in order to achieve promotion keywords in the search engine rankings,and next to reach a conversion, and this conversion is one condition to judge to the success of an e-commerce.

- Site structure

S.CN' s site uses a flat structure, all products are in the same level, the structure is clear that the naming methods is the URL for a brand shoes name. The navigation divided into man, woman, accessories, etc., the brand navigation is on the right side.

Fig. 2. The display for homepage and navigation of S.CN

There are 614 links in the homepage and inside page.After removed 200 no-follows (as shown in Figures 3), there are two entrances for search engines, one is the brand navigation, the other is the classification of goods (as shown in Figures 2). A lot of link which has no collection value has been removed at the website, it is much convenient for search engine to collect index page. Also S.CN makes a lot of brand project page for brand shoes.

A clear structure not only can effectively improve customer experience, but also have a positive response to viscosity for customer. In addition, site with complete structural also appears extremely "affinity" for search engines; this lets SEO activities more efficiently and quickly.

- Correlation links

User can easily find a good what they want to find because of correlation links which is good distribution on the homepage. For the filter condition, screening results page for no product tell the search engine no longer grabbing through the robots document this is a very good solution for the problem of site collection.

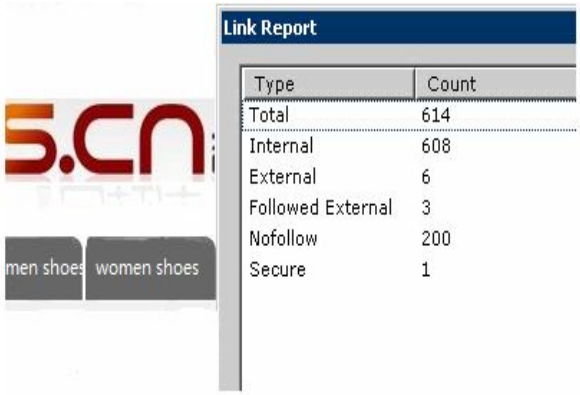


Fig. 3. The display for website links

Site map can be divided into two kinds, one is a HTML for user, and other is an XML for search engine. Both of them do well, such as: www.s.cn/page-sitemap.html--www.s.cn/sitemap.xml. The site of a sitemap file is divided into multiple files for the page which much good information on it, so that people can make search engine read and grab faster .

```
-<sitemapindex>

 -<sitemap>
  <loc>http://www/s.cn/sitemap1.xml</loc>
  <lastmod>2013-12-09</lastmod>
  </sitemap>
   -<sitemap>
  <loc>http://www/s.cn/sitemap2.xml</loc>
  <lastmod>2013-12-09</lastmod>
  </sitemap>
   -<sitemap>
  <loc>http://www/s.cn/sitemap3.xml</loc>
  <lastmod>2013-12-09</lastmod>
```

</sitemap>
-<sitemap>
<loc>http://www/s.cn/sitemap1.xml</loc>
<lastmod>2013-12-09</lastmod>
</sitemap>
</sitemapindex>

404 pages： Settings error page: according to manually input incorrect URL for the users, system will automatically jump to 404 error page, after a few seconds it will jump back to the home page.

- Setting path for URL

It is conducive to the search engines that all websites are static in the same level. All of the columns use brand words' namely semantic, such as: nanxie.html,nvxie.html, tongxie.html, fuzhuang.html, new-newnanxie.html, nike-brand.html, adidas-brand.html, lining-brand.html.

- Image optimization

All pictures for goods have Alt in order to increase the density of keyword. And it is one of sources for rate (as shown in Figures 4).

Fig. 4. The display for image optimization

- Code optimization

Homepage only list h4 for production, the improved method is that corrects h1 - h4 at the location of the page. Many pages only have a label—h1, at the same time, the product classification has no relevant tag. Some scripts are't called besides Style Sheets and JavaScript (as shown in Figures 5).

3.3 Summarize for S.CN

In general, optimization for S.CN is done better, especially the URL. There are some instructions in the most of pictures. And titles and descriptions are distributed well. In the term of code optimization, it does not make full use of tag H. Selection is a problem for most e-commerce site, but S.CN uses Robots to set some pages which have no search results forbidden search engines grabbing, it also removed many filter conditions through the no-follow.

The honor that S.CN is the three tops of Chinese footwear B2C enterprises is a show of success for S.CN. S.CN becomes the footwear e-commerce sites leader through the wave of electronic commerce and successful network marketing.

In addition, except SEO,S.CN also has more network marketing successes, such as T-mall flagship store, QQ mall flagship store, baidu, rebate sites, trust login, search engine marketing (SEM) (with SEO difference), mobile phone client (IOS, android), etc. This is the secret of success that these are the most fashionable means of network marketing. Of course, S.CN is insufficient, such as the disputes of genuine and fake goods.

4 The practice of the new methods of network marketing

The core of marketing is surprise and continuous innovation, to capture market opportunities from the conventional marketing environment. Actually, the network marketing is the same of the traditional marketing, so it is very important to innovate. The influence which network marketing brings to enterprise is unprecedented. With the development of Internet, the replicated marketing mode can't meet the needs of the consumers.

4.1 The marketing innovation which break through the traditional model

Due to breaking the traditional marketing and using the innovative marketing mode, the network marketing becomes so hot.Never copy others' success, and just to find the needs of users.

4.2 Product details for network marketing

At present, many website belongs to advertising type site which just has display for product and company introduction. Product features describe what the product is, but consumer will purchase the products after understanding its condition, not watching the features. The description for product should be related to customers need, saving money and time, increasing wealth, providing convenience, easy to know, avoiding the risk, etc. In addition, it should enrich the vivid data. While the name is very important, the description should manifest the value of products.

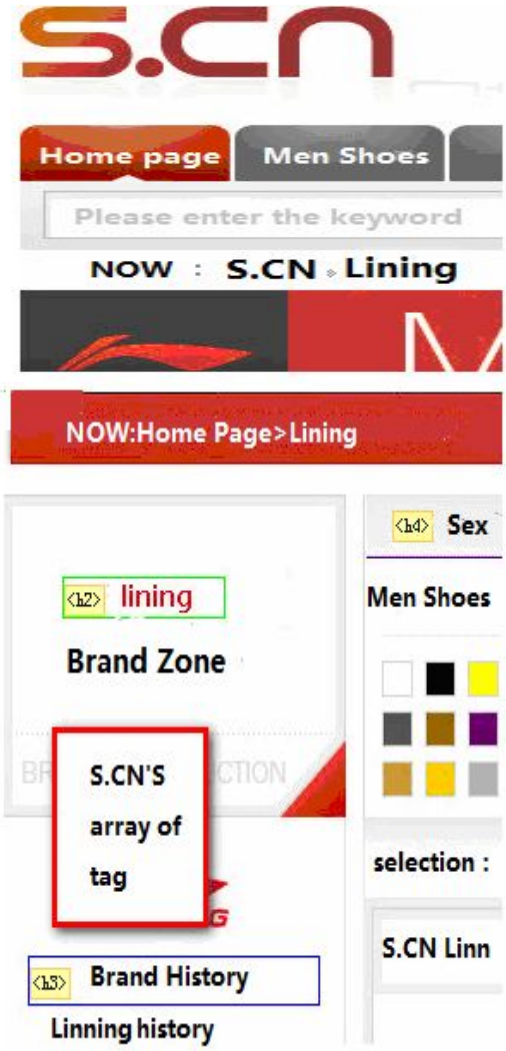


Fig. 5. The display for code optimization

4.3 Virtual collocation

Because of network marketing's virtual, it does not tally with the imagination or customer characteristics, etc., especially for clothes and shoes, so that customer satisfaction is low. Vendors can design software to make customers personalized data input and automatically matching, choose for customers to provide more accurate reference and determination, improve the credibility of the network marketing, the success rate and customer satisfaction.

4.4 Personalized design

Enterprise can provide personalized service for consumers can fully display their needs on the Internet and design them by themselves, product such as clothing, shoes, hats and bags and so on can use this way. Personalized design can meet the personalized needs of consumers greatly; improve the competitiveness of the enterprises and customer loyalty through the provision of special products, excellent quality, and high value services. To avoid the intermediate links, emphasizing the innovation design, after-sales service management and credibility of integration management efficiency can achieve rapid formation and development of fission of the market.

4.5 Cheers

Along with the network popularization, the network marketing tools and methods emerge in endlessly; Weibo should be the most popular methods in China. But in abroad, a new kind of network marketing method-- Cheers, is entering people's field of vision, its special is that you can share happiness and transfer positive energy combined with marketing at the same time, so that you can reduce the resistance of consumers. The largest business model of Cheers is the conversion of the fans and users which have advertising effects.

The market value of Cheers is that : the iPhone has no application that share to express love and gratitude, and Cheers just fills the blank of market; Although the profit pattern of Cheers also is not clear，we can know the potential of it, such as local store can attract users to focus on their service by Cheers, or use influential users on the brand promotion.

4.6 The innovative methods for network marketing by the development of network technology

Progress of network technology provides a new platform and technical support for the network marketing. At the same time, the innovation of network marketing has not only confined to the network, but also to the integration of online and offline, as well as the network marketing can provide more innovative ways.

Along with the development of mobile terminals and 3g network, Internet marketing represented by phone mobile also gradually expands. According to the report, it predicts that in the next five years there will be more users can via mobile devices rather than desktop access to the Internet and mobile Internet will surpass desktop development speed of the Internet.

Marketing innovation of the traditional enterprise is integrated by offline to online, while the network marketing began to emerge from online to offline.

4.7 The actual effect for S.CN by network marketing

S.CN became from an unknown web site to the top three of footwear e-commerce sites, which the effort and hard work is obvious to see, the source of its success is the correct market orientation and mature marketing concept. But at the same time, what the network marketing promote and influence cannot be ignored, and what is the embodiment of the network marketing is the most intuitive.

We can see the change of PR value obviously before and after the network marketing from Table 1, although the decrease in the number of total included, this is that the website optimize the integration.

TABLE I. THE STATISTICS FOR PR(PAGE RANK)

PR	Before optimization	After optimization
6	1	2
5	1	4
4	3	9
3	2	13
2	11	14
1	7	24
0	1120	557

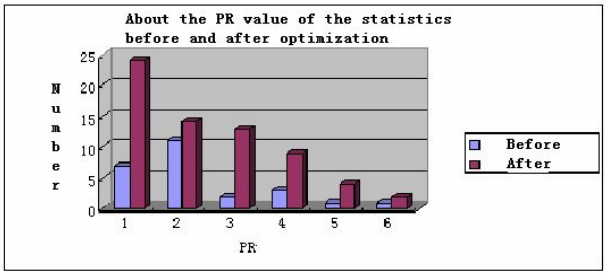


Fig. 6. The change of PR value (before and after optimization)

TABLE II. THE STATISTICS FOR BACK LINK

Back link	Before optimization	After optimization
http://www.s.cn	165	21
http://www.s.cn/s/hao123.html	50	24
http://www.s.cn/adidas-brand.html	25	20
http://www.s.cn/page-huodaofukuan.html	23	16
http://www.s.cn/page-abouts.html	18	13
http://www.s.cn/page-gouwubaozhang.html	17	3
http://www.s.cn/#r-208-865402-498-m.html	14	16
http://www.s.cn/#r-208-815688-499-m.html	9	1
http://www.s.cn/artlist-206.html	8	1
http://www.s.cn/message.html	8	1
http://www.s.cn/new-newlining.html	3	68

And the number of outside chain also has increased significantly, which also rely on the promotion of network marketing, the promotion makes the website more well-known and get more high-quality links. But it should be pointed out that the main domain name mostly is used to be the link name of S.CN, it is still insufficient for brand promotion. But after optimization, the problem is settled, as shown in Table 2. According to good distribution, the network marketing will be better put into effect.

5 Summarize

How to find the guiding ideology, innovation path, and the specific measures for the implementation of network marketing has been plagued by the marketing world, but we can find that the network marketing innovation is not just a simple traditional marketing network, just held on the fundamental change that the enterprise marketing activities to communicate with consumers combined with the network platform, the construction is different from the traditional new initiative, also enhance the accuracy and efficiency of marketing work between enterprise and customer.

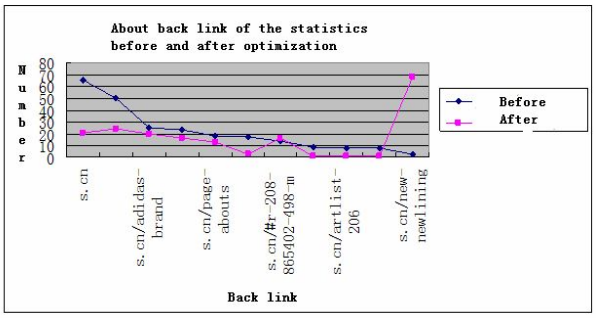


Fig. 7. The number of back link (before and after optimization)

It is obvious in this case—"S.CN" that some marketing which was successful have does not conformed to the current network marketing activities. With the concept of e-commerce matures, network marketing will also become familiar, therefore, attempts to "speculation" have become unrealistic. It will be successful just have the courage to use in the real case.

Network marketing is unalterable. Enterprises want to succeed only in combination with its own situation to make the most reasonable choice and be active to the changes in the Internet and new technologies and concepts. With learning attitude, you should have the courage to try new things, and apply to actual operation. You just be located in the indefectible position in the wave of network marketing which sum up the experience in actual combat, and timely adjust the network marketing strategy.

Acknowledgment

This paper funded by key discipline construction-Software Engineering(XXKZD1301) of Shanghai Second Polytechnic University.

References

[1] Zhikun Hu, Wenrong Jiang, Cuihong Wu, Hailan Pan, Jian Chen. The Methods and Analytical Studies of Taobao Stores Promotion [J]. The Ninth Wuhan International Conference on E-Business (WHICEB 2010), Wuhan, China: IEEE Press. 2010(5):118-122. ISI: BRR19.

[2] Wenrong Jiang, Shiwei Lin. Information Technology in the E-commerce Marketing Strategy of OAWAY Website [J]. 2013 3rd International Conference on Electronic Commerce, Web Application and Communication (ECWAC 2013), Wuhan, China: Advanced Materials Research (Volumn 859), 2014(1):551-555.

[3] Zhang yan: Ningbo small and Medium-sized Enterprise on –"luxury feast" for Informatization Application [J] Small and Medium-sized Enterprise on Study. 2008.6: 5-19.

[4] Yanlin Lv. The formation mechanism of service brand competitiveness research [D]. Shandong University, 2012.

[5] Fengjun Liu, Bing Lei, Yanxia Wang. The experience of consumer demand and marketing strategies for economy era [J]. China industrial economy, 2002(8):81-86.

[6] Hongli Guo. Study on the theory and method of customer experience management [D]. Shanghai Tongji University, 2006.

[7] Xiuying Wang. The analysis of the experience economy and value [J]. Journal of Beijing industry and commerce university: social science, 2005(1):46-49.

[8] Jingming Chen. Data mining and the customer satisfaction of e-commerce [J]. 2004 China (fuzhou) proceedings of international conference on electronic commerce and academic exchanges. Beijing: ZhongGuoKe association journal 2012:36-37.

E-Commerce in Saudi Arabia: Acceptance and Implementation Difficulties

E. Makki, and L.-C. Chang
Department of Electrical Engineering and Computer Science, School of Engineering, The Catholic University of America, Washington, DC, United States

Abstract - *Despite the fast growing economy, the dramatically increasing internet usage and the significant purchase power of Saudi Arabia compared to other countries in the Middle East, the evolution of E-Commerce remains slow and is inconsistent with such growth nor with the growth of E-Commerce in developed countries. This paper is conducted for exploration purposes to initially determine the most significant inhibitors of the acceptance and implementation of E-Commerce in Saudi Arabia for both merchants and individuals in order to conduct further quantitative research. A small sample of merchants and consumers is selected. The Saudi merchants, individuals and Saudis in USA data collected for this research is used to analyze inhibitors of acceptance and implementation of E-Commerce in Saudi Arabia. The most common inhibitors resulted from this research will be used for constructing surveys on a large nationwide sample for future studies. The study also proposes a framework for key decision makers and to be used in the further research in compliance with gathered data based on achieved analysis in this research.*

Keywords: E-Commerce, Saudi Arabia, Government Support, Payment Systems, Awareness

1 Introduction

E-Commerce adoption and implementation in developed and developing countries have been extensively researched with focus on consumer behavior or business issues. However, the amount of studies concentrating on E-Commerce in Saudi Arabia is inadequate [1], [2]. Furthermore, most of the studies either focus on businesses or consumers. This study will focus on E-Commerce implementation and adoption for both businesses and consumers.

As an important economic growth factor, E-Commerce became vital for many businesses to remain in competitive markets. Therefore, having an online option for sales is dramatically increasing in developing countries. However, this increase is not applied to E-Commerce in Saudi Arabia. E-Commerce in Saudi Arabia is trending slowly even with its strong economy and vastly growing internet usage [1], [3], [4].

This paper is attempting to study the inhibitors of E-Commerce acceptance and implementation in Saudi Arabia. In this study, we explore the issues facing consumers and retailers to prepare for further quantitative research of E-Commerce, and are trying to determine which inhibitors affect most and to what extent. We are also trying to explore unstudied factors to determine the rate of their effectiveness on acceptance and implementation of E-Commerce in Saudi Arabia.

2 Literature Review

Several broad researches have been conducted to study E-Commerce acceptance in developing countries. Such studies varies between focusing on general perspectives like social issues, and particular perspectives like online payment methods, trust, government role and delivery systems [1], [2], [5]–[8]. Some researches focus on large regions [9], and some focus on specific countries like Saudi Arabia [1], [5]–[7], [10], [11].

E-Commerce adoption factors and implementation issues in Saudi Arabia have been studied [11]. In this section, we provide a brief overview on the studies that concentrate on adoption factors, issues related to implementation, demographics of internet users, and proposed questions to answer for both consumer and business in development countries.

2.1 E-Commerce in Developed Countries

In the 90's, the evolution of E-Commerce practice have begun in most developing countries. The quick evolution of E-Commerce transactions in developed countries have been noticed [6]. Such practice is commonly acknowledged to be a factor of economic progression in developing countries [10]. In 2002, the global spending growth resulting from E-Commerce transaction reached around 0.27 trillion USD significantly jumped to 10 trillion after a decade [6].

2.2 E-Commerce in Saudi Arabia

Some research and statistical reports illustrate significant growth of internet, mobile and social media usage in Saudi Arabia. Most of these studies focused on either the business or the consumer perspective, but not both. Other researchers conducted several studies focusing on either business or consumer, as well as on both [2]-[4]. Additional researches studied the perspectives of business only regarding the acceptance of E-Commerce in Saudi Arabia [3], [5], [10], [12]. For examples, Al-Hudaif & Alkubeyyer studied internal and external aspects of E-Commerce acceptance [10].

Almousa conducted a research to profile online shoppers in Saudi Arabia [7]. An interesting report by de Kerros Boudkov Orloff discussed different factors related to business, consumers, government and banks [13]. Several solutions, like *The Five-Part Conceptual Model* framework proposed by AlGhamdi, Drew and Alhussain, among others are proposed. AlGhamdi's framework discusses critical issues that significantly affect the progress of implementing E-Commerce in Saudi Arabia. Those factors are: secure and trusted online payment methods, consumer protection, market place regulations, certification authority and delivery systems [6].

According to the report published by Sacha Orloff Consulting Group, computer and internet access for Saudis reached 65.8% of population [13]. Research and statistical reports on E-Commerce did not consider if transactions are associated with a local or an international vendor. The statistical report conducted by discoverdigitalarabia.com shows that results from respondents of a survey on E-Commerce in Saudi Arabia without indication if all or some transactions are from local vendors [14]. However, the report, as well as the report by Sacha Orloff Consulting Group, present valuable statistical results for researchers, especially for qualitative research, such as purchase frequency, monthly spending, payment method used, and first online purchase. In addition, both reports offer statistics about mobile usage in online purchase [13]–[15]. The usage of mobiles among Saudi population has outstandingly grown to 1.8, meaning for every 10 individuals there are 18 mobiles, to reach 95% of residents. Such results open doors for research to consider mobile usage for online purchase or at least for advertising the products.

Similarly, there is also remarkable growth in social media usage in Saudi Arabia. Statistical reports and information graphics about social media usage in Saudi Arabia puts it in top of other countries in the region and the world. According to The Social Clinic report early 2013 titled *The State of Social Media in Saudi Arabia 2012*, Saudi Arabia ranked number one in the world for the number of daily viewed videos, 90 million, on YouTube! [16]. Saudi Arabia also ranked number one in the world in Twitter growth rate of 3,000% which is 10 times the average global rate. Third, 2 out of 6 million, Facebook users in Saudi Arabia only used mobiles to access Facebook last year [16]. Research and reports about Saudi Arabia's growth in internet usage, social media and mobiles should consider such factors in their studies because some research results stated that internet and technology negatively inhibit the acceptance and implementation, which might not be accurate.

2.3 Factors influencing the evolution of E-Commerce in Saudi Arabia

There have been several overlaps between studies discussing inhibitors of E-Commerce in Saudi Arabia. Most agreed on a number of factors including but not limited to:

- Government regulations
- Delivery system
- Online payment
- Lack of trust
- Internet security and privacy
- Information and Communication Technology (ICT) infrastructure
- Business systems integration
- Banks and credit cards issuance
- User readiness and awareness
- Institutional roles

Those inhibitors have been studied using different approaches with a variation of data samples [3], [4], [6].

2.4 Review Conclusion

The inadequate research volume conducted to study E-Commerce in Saudi Arabia and the lack of large assortment statistical reports create a chance for further research [11]. Some common factors may be diminishing by other researchers like the user readiness and awareness, and ICT infrastructure. In addition, some factors could be added for further studies such as mobile usage and social media. More research related to mobile and social media could result in more reliable findings [13]. The most essential issues that researchers focus on are: payment methods, E-Commerce regulations and delivery systems. Although those issues are considerably obstructing E-Commerce in Saudi Arabia, they have not yet been resolved.

2.5 Questions of the Study

The objective of this paper is trying to answer the following questions in order to conduct further quantitative research on a large sample:

- What are the main differences of inhibitors in accepting E-Commerce in Saudi Arabia between local Saudi consumers and Saudis in USA?
- Do mobile usage and social networking play any role in E-Commerce in Saudi Arabia?
- What are the most common aspects that negatively or positively affect the acceptance and implementation of E-Commerce in Saudi Arabia?
- What solutions are most likely to be successful to resolve the problems of acceptance and implementation of E-Commerce in Saudi Arabia?
- What solutions can be added to escalate the rate of acceptance and implementation of E-Commerce in Saudi Arabia?

3 Methodology

A qualitative approach for this study is utilized for primary exploration research. The qualitative approach helps gain profound understanding of the inhibitors of E-Commerce in Saudi Arabia. The data collected for this study is of multiple sources and targets. The targets are: local Saudi internet users, Saudi consumers in USA and local Saudi retailers. Due to the lack of reach to Saudi retailers and local

consumers, data from a recent study is used [6]. Furthermore, 10 Saudi citizens in the United States are interviewed with similar consumer questions from the previous research along with additional 2 questions related to mobile E-Commerce usage and social media involvement in their E-Commerce practice. Table 1, 2 and 3 show the questions and answer options for local Saudi consumers, local retails and Saudis in USA consumers respectively.

Table 1: Questions and answer options for local Saudi consumers [6]

"What factors inhibit or discourage you from buying online from e-retailers in Saudi Arabia?"	**"What would enable you to buy online from e-retailers in Saudi Arabia?"**
'Lack of physical inspection'	'Physical shop as well as online shop'
'Lack of clear regulations and legislation'	'Government supervision and control'
'Lack of online purchase experience'	'Competitive prices'
'Don't trust that personal info will remain private'	'Trustworthy and secure online payment options'
'Do not know e-retailers in Saudi Arabia'	'Owning a house mailbox'
'E-retailers are not trusted in terms of mailing products in quality same as specified'	'Well-designed retailer websites'
'Lack of mailbox for home'	'Easy access and fast Internet'
'Uncomfortable paying online using credit cards'	'Provision of educational programs'
'Lack of English language understanding'	'Local banks make owning credit cards easier'
'Difficult access to the Internet and slow speed'	

4 Results and Discussion

The results from AlGhamdi, Drew and Alhussain research show both anticipated and unanticipated results [6]. Low percentage in the results could indicate more findings. For example, the percentage to the discouraging factor for retailers of *"Delivery issues"* is unexpectedly low (10.8%) whereas many researchers concluded that this factor is a crucial inhibitor for retailers [8], [10]. However, this result is consistent with researches from consumer perspective [13] as delivery systems are not major obstacle to retailers as it is to consumers. Table 4 shows the results for consumers indicating that the top inhibitor for consumers to accept E-Commerce is physical inspection (53.8% as an inhibitor and 62.4% as a motive) slightly higher than the expected government regulations (52.4% as an inhibitor and 52.6% as a motive). This might be due to the culture of Saudis who prefer to physically inspect the item before purchase to avoid buying defected or wrong item and/or the lack of trust to sellers.

Table 2: Questions and answer options for local Saudi retailers [6]

"What factors inhibit or discourage your company from implementing an online system to sell on the Internet?"	**"What factors help or encourage your company to implement an online system to sell on the Internet?"**
'Habit/Culture of people to buy is not encouraging'	'Trustworthy and secure online payment options'
'Lack of clear legislations and rules of e-commerce in the KSA'	'Government support and assistance for e-commerce'
'Lack of e-commerce experience'	'Develop strong ICT infrastructure'
'Products are not suitable to be sold online'	'Educational programs and building awareness on e-commerce'
'Poor ICT infrastructure'	'Provision of sample e-commerce software to trial'
'Lack of online payment options help to build trust'	
'Resistance to change'	
'We do not trust online sales'	
'Delivery issues'	
'Setup cost'	
'Cannot offer competitive advantage'	
'Not profitable'	

Table 3: Additional questions and answer options for Saudis in USA consumers

How often do you use your mobile to purchase online?	Does social media encourages your decision to buy online?
Always	Definitely yes
Most of the time	Probably yes
Sometimes	Uncertain
Rarely	Probably no
Never	Definitely no

Table 4: Results for positive and negative factors for consumers [6]

Inhibitors	%	Motives	%
'Lack of physical inspection'	53.8	'Physical shop as well as online shop'	62.4
'Lack of clear regulations and legislation'	52.4	'Government supervision and control'	52.6
'Lack of online purchase experience'	44.7	'Competitive prices'	50.0
'Don't trust that personal info will remain private'	41.0	'Trustworthy and secure online payment options'	41.5
'Do not know e-retailers in Saudi Arabia'	36.6	'Owning a house mailbox'	36.2
'E-retailers are not trusted in terms of mailing products in quality same as specified'	33.4	'Well-designed retailer websites'	33.8
'Lack of mailbox for home'	31.8	'Easy access and fast Internet'	32.6
'Uncomfortable paying online using credit cards'	26.3	'Provision of educational programs'	29.0
'Lack of English language understanding'	21.9	'Local banks make owning credit cards easier'	20.6
'Difficult access to the Internet and slow speed'	7.5		

Price also has a role to enable adoption of E-Commerce in Saudi Arabia as it is ranked third with 50%. This should be higher compared to other countries where online prices are likely to be lower than store prices as a general knowledge. Perhaps the reason is the cost of E-Commerce to vendors in Saudi Arabia is higher [11]. Top consumer inhibitors that can be enabled by awareness efforts are related to the culture of consumers like the necessity of physically inspecting items, getting used to shop online and exposure of online retailers. Social media could play a significant role in this aspect since the penetration and usage is high [13], [14], [16]. By involving social media, companies, governments and other consumers could help increase awareness, diminish security and privacy fear and collaborate in proposing solutions for their issues. Moreover, trust issues facing vendors could be reduced by engaging with customers directly through social media. Table 5 shows results for positive and negative factors for retailers.

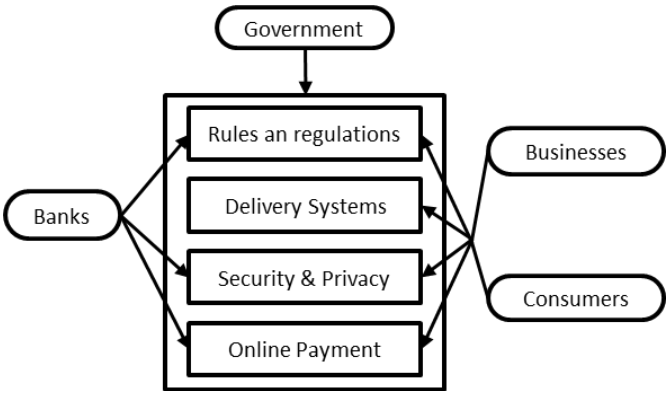
Table 5: Results for positive and negative factors for retailers [6]

Inhibitors	%	Motives	%
'Habit/Culture of people to buy is not encouraging'	42.6	'Trustworthy and secure online payment options'	58.1
'Lack of clear legislations and rules of e-commerce in the KSA'	35.8	'Government support and assistance for e-commerce'	53.1
'Lack of e-commerce experience'	35.8	'Develop strong ICT infrastructure'	39.9
'Products are not suitable to be sold online'	25.0	'Educational programs and building awareness on e-commerce'	31.1
'Poor ICT infrastructure'	22.4	'Provision of sample e-commerce software to trial'	25.7
'Lack of online payment options help to build trust'	20.9		
'Resistance to change'	16.9		
'We do not trust online sales'	12.8		
'Delivery issues'	10.8		
'Setup cost'	10.1		
'Cannot offer competitive advantage'	7.4		
'Not profitable'	8.1		

Online payment methods is ranked 6th with nearly 21% as an inhibitor for retailers, but it is a strongly requested motivator, ranked 1st with about 58% for retailers to adopt E-Commerce. This complies with other research that online payment is an extremely significant issue for both retailers and consumers [3]–[6], [10], [13]. Furthermore, government regulation of E-Commerce is another substantial issue, not only for retailers

and consumers, but for banks and perhaps delivery systems as well. Culture of Saudis insisting on physical inspection ranked as the top inhibitor for retailers, which underlines the essentiality of awareness. Perhaps the involvement of government and institutions would increase such consciousness [4], [13].

For exploration purposes, a decision was made to minimize the inhibitors in order to conduct further research on most confirmed major factors for each perspective. For consumers, it is found that online payment methods, delivery system, trust (security, privacy and vendor trust) and lack of regulations are the primary inhibitors to accept E-Commerce in Saudi Arabia [6], [8]. As for retailers, online payment methods, security & privacy, lack of regulations and E-Commerce systems integration are the most discouraging factors. System integration involves financial, inventory and banking systems. The lack for skilled ICT professionals also plays a role. Obviously, there are shared inhibitors for both, which is related to third parties like government and banks. Figure 1 illustrates the relationship of inhibitors and key players.



Figuer 1: Relationships between inhibitors and key players

For responses gathered from Saudis in USA, it is assumed that the only overlap is in security and privacy to a limited extent. This is because they now live in a developed country and issues such as payment methods and delivery systems are minimal. However, the added questions related to mobile usage and social media provided interesting results. Table 6 shows results we surveyed for mobile and social media.

Table 6: Results for mobile and social media

How often do you use your mobile to purchase online?	%	Does social media encourages your decision to buy online?	%
Always	0	Definitely yes	60
Most of the time	10	Probably yes	20
Sometimes	60	Uncertain	10
Rarely	20	Probably no	00
Never	10	Definitely no	10

Many Saudis in USA are encouraged by social media to buy online. While purchasing online using mobile is moderate. Since the culture, despite of the environment, of Saudis in USA is similar with local Saudis, the results could encourage retailers to utilize mobiles and social media as channels of marketing as well as increase awareness of E-Commerce. Government and institutions could also use social media to publicly increase awareness.

5 Suggested Framework

Based on the results of this study and previous works, we conclude that the main inhibitors of accepting and implementing E-Commerce in Saudi Arabia that affects both retailers and consumers are:

- Secure Online Payment Methods
- Government regulations
- Delivery Systems
- E-Commerce Awareness

5.1 Secure Online Payment Methods

It is obvious that credit cards are the most convenient method of payment for conduction online transactions especially in developed countries. It is easy to obtain and preferred by both merchants and consumers due to ease of use and ability to track transactions. However, there are many inhibitors in Saudi Arabia preventing individuals from seeking or obtaining credit cards. Some of the main factors are cultural and/or religious reasons, the concerns of privacy and security, the eligibility to have a card issued from banks and avoiding paying fees and interests. Other secure payment methods do exist in Saudi Arabia such as PayPal and SADAD. PayPal is a globally recognized secure online payment method, while SADAD is a local secure payment method targeting large companies for bill payments. PayPal can be a mediator between customers, banks and retailers with excellent features such as consumer protections. PayPal holds merchant's payments for a period of time in case there are complaints. It also facilitates the ability to refund consumers in case they are not satisfied. A problem with PayPal in Saudi Arabia is that it only accepts credit cards or wired deposit to PayPal balance and not connecting local bank's account.

Saudis tend to avoid the use of credit cards due to the interests, which is forbidden by Islam. There have been several attempts by banks to offer Islamic credit cards. However, such attempts are not sufficient for broader use in E-Commerce. The reasons behind that are the high fees and requirements. Most banks require minimum amounts to be deposited as well as a regular monthly direct deposit, like salaries, in order to get a credit card. Yet, a research by Almousa stated that the majority of internet users in Saudi Arabia who are ready to buy online, more experienced and spend more time online, are residents aged between 18-35 who are either dependents or students [7]. This segment of the Saudi internet users are the most appealing to buy online and should not only be targeted by retailers, but also by banks as well. Most of them prefer prepaid credit or debit cards as the best solutions for online payment. Unfortunately, only a few

banks offer such cards with very poor customer service and some of them with hard to achieve requirements. Banks should target young Saudis by offering them easy to get prepaid credit cards. In addition to prepaid cards that can be linked to PayPal, banks should also work on allowing their customers to connect their PayPal accounts to their debit accounts. Another issue with banks is that they are not cooperative in disputing charges made on their cards. Consumers must file a complaint and pay a fee with a long waiting period to get their money back. This could be one of the inhibitors for some consumers to use their prepaid/credit card for online purchases. Up to date, banks offer online merchant accounts with unfeasible rates, which discourage retailers to use local payment systems with regular ATM cards. SADAD is a billing system connected to local banks and large vendors like utility providers. The list of providers is increasing, but growth is not complying with the demand for E-Commerce growth. In addition, some companies complain that the fees are not feasible to participate. Unlike PayPal, SADAD is not offered to individuals to accept online payments.

5.2 Government regulations

The majority of research conducted to study E-Commerce in Saudi Arabia share similar results that the lack of government regulations is vital for the acceptance and implementation of E-Commerce in Saudi Arabia. There are no E-Commerce rules or regulations in Saudi Arabia to date. However, there is a government consumer protection agency that is not actively supportive. Consumers do not trust that this agency can protect them if they complain regarding offline transactions. Therefore, they cannot trust the agency for online protection. The significance of having such regulations for both retailers and consumer is crucial for conflict resolutions, online business practice and consumer protection for transactions conducted online. Additionally, new regulations for individual online business licensing are needed since current rules insist on renting an office in a merchant location. Many individuals try to sell self-made items, like clothes or art works, online using a website and accepting orders by phone calls or emails. They wait for direct deposit payments in scanned proof and contact the buyer to meet in person for delivery. Such individuals face trust issues because they are not licensed. Buyers are afraid that the online seller is not trustworthy or there is no place to file complaints if necessary.

Since the government plays a key role in online privacy and security regulations, both consumers and retailers are affected by regulations. Retailers should secure their transactions and consumers also need to secure their computers. The government involvement in online security and privacy is also very weak. Victims of identity theft struggle to complain and to resolve their issues. The government should establish E-Commerce rules and regulations for E-Commerce transactions as well as licensing for individuals. It should also play a more productive role in resolving both consumer and retailer issues. Government agencies should be more involved in helping researchers to study this aspect by providing statistical information to assist in quantitative research.

5.3 Delivery Systems

The delivery system in Saudi Arabia mainly depends on P.O. Boxes rather than on-street addresses. In Saudi Arabia, mail is mostly delivered to mail boxes in post offices. Delivery issues are not just related to E-Commerce, but also to other services like pizza delivery. The main way to get items delivered to your door is to instruct drivers with landmarks. A physical address system called "Wasel" is launched in 2005 by the local Saudi Post, which is government owned, to set physical addresses. The service is growing slowly because individuals must request it first by themselves. Moreover, only a few major cities is covered and limited subscribers have been adopted. Other commercial delivery systems in Saudi Arabia, like FedEx, UPS, DHL and an Arabic company called Aramex, exist and are used in local and international delivery. Unlike the way those companies deliver in other countries, the customer must provide a phone number to be contacted for delivery or pick up. The Saudi Postal Services and local commercial delivery companies should adopt a system that can be easily integrated into merchants' systems for goods delivery. For now, a temporary solution employing Google geocoding and GPS could be sufficient. Consumers could click on the map in merchant's website to set their home locations and it will be added to the purchase form. Retailer can then use GPS, or Google Maps App, to drive to the location. This might not be perfect, but needs a closer investigation by researchers. An online service called enwani.com provides such solution [17].

5.4 E-Commerce Awareness

Since one of the major inhibitors for both consumers and retailers is the need for physical inspection of items as a cultural factor, awareness of the benefits of E-Commerce is crucial. Before promoting E-Commerce, its benefits should be appealing to consumers. Item showcasing, website and purchasing process user experience, prices and delivery should be good enough for online consumers to consider. Accordingly, merchants should put more effort on creating clear item description with various images and clear specifications on their websites. They should also eliminate other costs, like marketing, store rent and manpower to reduce the prices. This could be done by starting with providing online only goods, which should not have extra costs like store, and set lower prices to test consumer acceptance. Price could be further lowered by offering warehouse pick up with physical inspection to increase the adoption to buy. To increase the awareness, social media can be utilized as it is significantly growing in Saudi Arabia and users might help sharing E-Commerce awareness campaigns [13], [14], [16]. According to a report by Yazeed Altaweel published on Discover Digital Arabia website, 64% of online stores use social media to promote their products and services [18]. Such campaigns should be collaborative by government, businesses and even individuals to spread the word. In addition to social

media, other trusted media sources like TV and street signs could be helpful. Not only convenience should be promoted, economic and cultural benefits should be explained to increase public participation. The success of E-Commerce awareness rely on other factors like easy payment methods, clear regulations, better delivery systems and the extent of business participation, where government can play a significant role.

6 Conclusions

This paper is conducted to determine the most significant inhibitors of the acceptance and implementation of E-Commerce in Saudi Arabia for both consumers and retailers in order to conduct further qualitative research. The study also proposes a framework that can be used in further research in compliance to data collected based on achieved analysis in this research. The most common inhibitors resulted from our study are: (1) Deficiency of government involvement with rules and regulations governing E-Commerce transactions, consumer protection, online business licensing and resolving security and privacy issues. (2) Secure online payment methods that both retailers and consumers can rely on in a way that increases the acceptance and implementation of E-Commerce. (3) Delivery systems that is not applicable to encourage the delivery of goods from retailers to consumers. (4) E-Commerce awareness of the economic and cultural benefits to the country.

Major implications of the study are the factors that affect both consumers and retailers which are related to the same key players. E-Commerce government rules and regulations apparently are the most significant inhibitors to E-Commerce in Saudi Arabia to consumers, retailers and sometimes to banks. Similarly, secure online payment methods affect both consumers and retailers. Delivery systems, on the other hand, affect consumers more than retailers. A shared factor by all key players is the E-Commerce awareness, which needs to be improved by all parties. Another implication consideration is the role of mobile and social media in E-Commerce. However, the role of social media is more adopted in encouraging online shopping than mobiles. In addition, social media, as proposed in this research, could significantly assist in increasing E-Commerce awareness in Saudi Arabia.

In conclusion, the findings of this paper complies with other works when investigate the most influencing inhibitors and motivators of accepting and implementing E-Commerce in Saudi Arabia. In addition, this study proposes other open opportunities for further research.

7 References

[1] M. Almousa, "Barriers to E-Commerce Adoption: Consumers' Perspectives from a Developing Country," *iBusiness*, vol. 5, no. 2, pp. 65–71, 2013.

[2] R. Alghamdi, A. Nguyen, and V. Jones, "A Study of Influential Factors in the Adoption and Diffusion of B2C E-Commerce," *Int. J. Adv. Comput. Sci. Appl.*, vol. 4, no. 1, pp. 89–94, 2013.

[3] R. AlGhamdi, S. Drew, and W. Al-Ghaith, "Factors Influencing E-Commerce Adoption By Retailers In Saudi Arabia: A Qualitative Analysis," *Electron. J. Inf. Syst. Dev. Ctries.*, vol. 47, no. 11, pp. 1–23, 2011.

[4] R. AlGhamdi, A. Nguyen, J. Nguyen, and S. Drew, "Factors Influencing E-Commerce Adoption By Retailers In Saudi Arabia : A Quantitative," *Int. Conf. Internet Stud.*, vol. 8, no. 10, pp. 83–100, 2012.

[5] S. M. Sait, K. M. Al-tawil, and S. A. Hussain, "E-Commerce In Saudi Arabia: Adoption And Perspectives," *A*, vol. 12, no. 1, pp. 54–74, 2004.

[6] R. AlGhamdi, S. Drew, and T. Alhussain, "A Conceptual Framework for the Promotion of Trusted Online Retailing Environment in Saudi Arabia," *Int. J. Bus. Manag.*, vol. 7, no. 5, pp. 140–150, Feb. 2012.

[7] M. Almosa, "Profiling E-buyers in Saudi Arabia : Saudi Arabia as A Promising Developing Country," *CROSS-CULTURAL Commun.*, vol. 7, no. 2, pp. 92–100, 2011.

[8] E. Aleid, "Delivery Issues in E-Commerce : The Case of Saudi Arabia" Ph.D. dissertation, University of Waikato, 2012.

[9] A. M. Aladwani, "Key Internet characteristics and e-commerce issues in Arab countries," *Inf. Technol. People*, vol. 16, no. 1, pp. 9–20, 2003.

[10] S. a Al-Hudhaif and A. Alkubeyyer, "E-Commerce Adoption Factors in Saudi Arabia," *Int. J. Bus. Manag.*, vol. 6, no. 9, pp. 122–133, Sep. 2011.

[11] I. Ahmad and A. M. Agrawal, "An Empirical Study of Problems in Implementation of Electronic Commerce in Kingdom of Saudi Arabia," *Int. J. Bus. Manag.*, vol. 7, no. 15, pp. 70–80, Jul. 2012.

[12] S. A. Al-Somali, R. Gholami, and B. Clegg, "An investigation into the acceptance of online banking in Saudi Arabia," *Technovation*, vol. 29, no. 2, pp. 130–141, Feb. 2009.

[13] A. de Kerros Boudkov Orloff, "E-Commerce in Saudi Arabia : Driving the evolution , adaption and growth of e-commerce in the retail industry," *Sacha Orloff Consult. Gr.*, no. June, 2012.

[14] "E-commerce Landscape in Saudi Arabia," *Discover Digital Arabia*, 2012. Internet: http://www.ddarabia.com/infograph/e-commerce-landscape-in-saudi-arabia/ [Jul. 22, 2013].

[15] K. Banga, "KSA e-Commerce transactions reach SAR 12 billion." Internet: http://www.smeadvisor.com/2010/06/ksa-e-commerce-transactions-reach-sar-12-billion [Dec. 20, 2013].

[16] "The State of Social Media in Saudi Arabia 2012," *The Social Clinic*, 2013. Internet: http://www.thesocialclinic.com/the-state-of-social-media-in-saudi-arabia-2012-2/ [Jun. 10, 2013].

[17] "enwani - home delivery made easy." Internet: https://enwani.com/ [Apr. 15, 2014].

[18] Y. Altaweel, "E-Stores In KSA," *Discover Digital Arabia*. Internet: http://www.ddarabia.com/infograph/e-stores-in-ksa [Nov. 11, 2013].

Exploring Mobile Computing Trends in Saudi Arabia: An Initial Investigation of M-Government and M-Business

Mutlaq B. Alotaibi
Vice Dean for Academic Affairs, College of Computer & Information Sciences
Al Imam Mohammad Ibn Saud Islamic University (IMSIU)
Riyadh, Saudi Arabia
motaibi@imamu.edu.sa

***Abstract*— The main purpose of this research was to gain an understanding of the adoption behavior of mobile computing in Saudi Arabia. In particular, it aimed to acquire new insight into mobile computing trends in Saudi Arabia, in order to develop hypotheses and formulate precise criteria for mobile computing evaluation. In order to achieve these aims, a focus group was created through the recruiting of eight participants with solid background knowledge of usability engineering and mobile computing. The focus group proposed a four-phase process: determination, qualification, categorization and evaluation of the mobile computing applications developed by Saudi organizations. During the determination phase, a total of two hundred and twenty seven (n=227) mobile applications, developed by organizations in Saudi Arabia, were determined. During the qualification phase, one hundred and forty two (n=142) mobile applications were qualified. During the categorization phase, only the qualified applications were categorized by the experts into a two-level categorization hierarchy. Finally, in the evaluation phase, the qualified applications were evaluated in terms of purpose, platform, visual appearance, content, organization and usability. The results herein revealed that 43% of the mobile applications in Saudi Arabia were M-Government applications, while 57% were M-Business applications. In addition, the study proposed a sample of thirty six (n=36) applications, representative of the population of mobile applications in Saudi Arabia, to have statistical significance.**

Keywords– mobile computing; Saudi Arabia; exploratory; mobile business; mobile government

I. INTRODUCTION

Mobile computing is a rapidly growing market, to illustrate the annual number of smart phones shipped to consumers is greater than the combined sales of automobiles and PCs [1]. In Saudi Arabia, the number of mobile phone subscriptions reached 53 million in 2013, representing a population penetration rate of 181.6% [2]. In fact, the widespread adoption of mobile computing has motivated the Saudi government and business organizations to deliver online services, through the use of smart phones, which allows users to install applications on their mobile operating systems and facilitates seamless and instant access to data anytime/anywhere [3]. Although the potential of mobile computing is well recognized, few studies have investigated these trends and key attributes in Saudi Arabia. Consequently, a gap in the literature has been identified.

This study attempts to overcome these gaps by exploring the status and trends of mobile computing in Saudi Arabia. It will contribute to the literature on Information Systems (IS) by exploring dominant characteristics of mobile applications and providing insight into key trends in mobile computing. In particular, it aims to formulate hypotheses to direct further research in the field by showing the differences between mobile computing types (M-Government and M-Business). It also aims to propose a sample of mobile applications in Saudi Arabia which truly represents the population, with statistical significance, in order to facilitate sample selection in future mobile computing studies. In addition, the study will provide a two-level categorization hierarchy of mobile applications in Saudi Arabia, along with application ranking and evaluation.

The remainder of paper will be organized as follows. Section two will present information on related studies and an exploratory study will be presented in section three. Section four will identify the results and section five will provide discussions as well as the implications of this research. Finally, conclusions will be provided in section six.

II. RELATED STUDIES

The increased research and practice in mobile computing demonstrates that this field is of significant importance. Over the past years, several studies have investigated the evaluation of mobile applications by drawing different conclusions. For example, Po et al. (2004) addressed the issue of usability evaluation for mobile technologies using heuristic evaluation, they presented contextual influences to bridge the realism gap [4]. In particular, the study proposed that the usability evaluators should be considered when evaluating mobile applications in order to control the variation in the context interpretation. A study by Kjeldskov and Stage (2004) proposed six techniques for mobile computing evaluation in controlled environments with regard to physical activities; they found that movement and navigation, while using mobile applications, could cause several usability issues, like missing a button in the interface [5]. Another study by Zhang and Abidat (2005) proposed a generic framework for the usability testing of mobile applications according to the nature of the

application; several challenges were identified with regard to mobile context, multimodality, connectivity, small screen size, different display resolutions, limited processing capabilities and power, and restrictive data entry methods [6]. In addition, Kaikkonen et al. (2005) examined the differences between usability studies of mobile applications in two environments: laboratory and field evaluations; they found that the same usability issues were identified in both environments [7].

More recently, a study by Verkasalo (2010), on the adoption of mobile computing, found that the motivation to use mobile applications depends on the application type, such as utilitarian motivations which drive the usage of mobile internet and map applications [3]. Franke and Weise (2011) proposed a framework of mobile computing quality assurance which provided key qualities of mobile applications, patterns of mobile applications development and metrics for mobile application testing [8]. A study by Alotaibi (2013) investigated the mobile service acceptance in Saudi Arabia, he found that Saudi was socially influenced in its adoption of mobile services [9]. The study assessed a mobile application in the financial services (entitled M-Tadawul), by revising the model of the unified theory of acceptance and use of technology (UTAUT); the results recommended further assessment of mobile applications in Saudi Arabia because mobile applications have not been studied in depth in the country and no hypotheses have been formulate, to date, for best use.

III. EXPLORATORY STUDY

In order to explore the mobile computing trends, a focus group of eight mobile computing specialists was created. The focus group included four men and four women in order to control the effect of gender. The group was also representative of four different levels of skill and knowledge: bachelor students, master students, PhD candidates and mobile computing practitioners. For each knowledge level, there was one male and one female working independently.

Initially, to generate a starting point, a brainstorming technique was utilized to identify the requirements for the exploratory study. The focus group proposed a process of mobile computing assessment which consisted of four main phases: determination, qualification, categorization and evaluation. The determination phase aimed to collect as many mobile computing practices as possible by searching application markets, such as App Stores and Google play. In the qualification phase, each mobile application was assessed by the focus group to determine whether it was qualified for further evaluation. The target differentiated between mobile applications that were developed by organizations or by individuals. In the categorization phase, the qualified mobile applications were categorized based on common attributes, such as sector, main activities, etc. This phase resulted in a two-level categorization hierarchy of the mobile applications. During the evaluation phase, the qualified applications were individually assessed by the focus group members. In brief, a four-phase process was followed to explore the level of adoption, trends and main attributes of mobile computing.

While implementing the proposed process the initial search of the application markets revealed thousands of irrelevant applications which were retrieved during the determination phase. Therefore, the search for mobile applications in Saudi Arabia utilized official indexes provided by Saudi websites, such as E-Government index [10], stock market index [11] and the Ministry of Commerce index [12]. The focus group members searched all of the mobile application stores, they noticed that some of the applications appeared in the different stores under different names. The output of the determination phase identified a total of two hundred and twenty seven (n=227) mobile applications that had been developed by organizations in Saudi Arabia. During the qualification phase, several applications were found to be: lower than the desired levels, duplicated versions or developed by individuals. The qualification phase identified a total of one hundred and forty two (n=142) mobile applications. Based on the sector, the categorization phase divided these applications into two types: M-Government and M-Business. Based on their main activities, the applications of each type (category) were then categorized into eight (n=8) sub-categories resulting in a total of sixteen sub-category (n=16).

TABLE I. THE CATEGORIZATION HIERARCHY WITH THE SUB-CATEGORIES RANKED, BASED ON THE NUMBER OF MOBILE APPLICATIONS IN EACH ONE (N) AND THE PERCENT

Category	#	Sub-Category	Rank	Applications (N)	Percent (%)
M-Government	**1**	Ministries	10	7	4.9%
	2	Education	8	9	6.3%
	3	Authorities	11	4	2.8%
	4	Health	12	4	2.8%
	5	Corporations	15	3	2.1%
	6	Other agencies	4	12	8.5%
	7	Government services	5	11	7.7%
	8	Government content	6	11	7.7%
M-Business	**9**	Banks and financial services	1	18	12.7%
	10	Energy and utilities	16	3	2.1%
	11	Retail and commerce	9	8	5.6%
	12	ICT	13	4	2.8%
	13	Media and publishing	3	16	11.3%
	14	Other businesses	7	11	7.7%
	15	Business content	2	17	12.0%
	16	Business services	14	4	2.8%
Total				142	100%

The output of the categorization phase resulted in a two-level categorization hierarchy, see Table 1, which demonstrates that the sub-categories were ranked based on the number of mobile applications in each one (N), and their percent.

The evaluation of these mobile applications utilized an instrument that was devised specifically for this study and which considered the following three parts: purpose, platform and evaluation. The aim of the first part was to determine the purpose of the mobile application as being one of three multiple choice answers: advertising only, displaying content and transaction processing. In the second part, all of the platforms in which the mobile application was running were identified using multiple choice answers for all of the listed mobile computing platforms, such as: iOS, Android and other OS. In the third part, each member of the focus group was instructed to independently rate the mobile application in terms of its visual appearance, content, organization and usability, using a ten-point scale ranging from very weak (1) to excellent (10). Conflicting rating results were then resolved through the use of short meetings and discussion panels which were supervised by the author.

IV. RESULTS

Figure 1 shows the mobile computing trends in Saudi Arabia, in terms of type, purpose, platform and evaluation. At a glance, it can be noticed that M-Business, content display and iOS were the most dominant trends with a service quality level that was more than 74% in terms of the application's visual appearance, organization, usability and content. In terms of the type of mobile computing, it was found that Saudi businesses were more progressive in their mobile computing than the government, with a score of 57% for M-Business which is 15% higher than that for M-Government. In terms of mobile computing adoption, it can be seen that all of the organizations, both public and private, utilized mobile computing to displaying content, compared to 44% of the organizations which utilized mobile computing for transaction processing and 17.6% for advertising only. With regard to platform trends, iOS was the most dominant OS in mobile computing, with 95.1% of the mobile applications running on iOS, compared to 59.9% running on Android and 17.6 running on other OS, such as Windows phone and Blackberry. In terms of the evaluation of the quality of mobile computing services, all of the evaluation factors were rated above 74% which is a very good indicator of the mobile application's usability and visual appearance as well as of its content and organization. The overall score for mobile computing was consistent with the scores for the four evaluation factors which reflected a balanced implementation of the mobile applications.

V. DISCUSSION

The purpose of this research was to formulate hypotheses, it is therefore important to identify the elements of mobile computing which are more progressive. The number of applications in M-Business was higher than in M-Government, which implies that M-Business is more progressive. However, this indicator cannot be trusted alone, as there are more private organizations than government ones. Therefore, it is important to consider a representative sample of mobile computing applications. With the aid of data mining tools, labeled SPSS Clementine, a sample representing all of the mobile computing applications was generated using the K-Mean clustering approach. The aim was to identify the quality levels and a set of applications that could serve as a prototype of each quality cluster. This resulted in a sample of thirty six (n=36) mobile applications, which were equally representative of M-Government and M-Business. This sample is shown in Table 2.

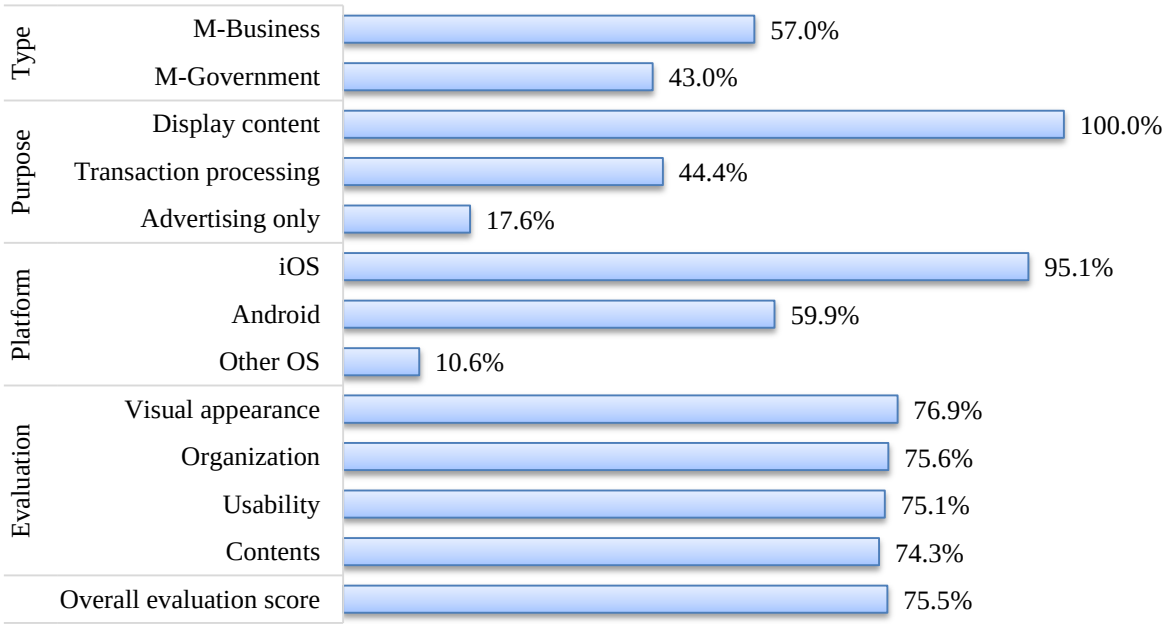


Figure 1. Mobile computing trends in Saudi Arabia

TABLE II. THE SAMPLE OF THIRTY SIX (N=36) MOBILE APPLICATIONS, WHICH WERE EQUALLY REPRESENTATIVE OF M-GOVERNMENT AND M-BUSINESS

Category	#	Name	Rank	Score
M-Government	**1**	Ministry of Commerce and Industry	55	79.1%
	2	Ministry of Foreign Affairs	30	83.1%
	3	Ministry of Health	15	89.1%
	4	Ministry of Higher Education	98	72.8%
	5	Ministry of Labor (MOL)	91	73.4%
	6	King Fahd University of Petroleum & Minerals (KFUPM)	130	59.4%
	7	Taibah University	110	69.4%
	8	Saudi Food & Drug Authority (Saudi FDA)	111	68.4%
	9	Saudi Arabian General investment Authority (SAGIA)	125	62.2%
	10	King Fahad Medical City (KFMC)	47	80.3%
	11	Saudi Post Corporation	121	63.4%
	12	Saline Water Conversion Corporation (SWWCC)	23	85.3%
	13	National Information center (NIC) at MOI	78	75.3%
	14	Riyadh Chamber of Commerce and Industry	69	76.6%
	15	Saudi Arabian Airlines (Saudia)	92	73.4%
	16	Saudi e-Government (Yesser)	12	90.0%
	17	General Directorate of Civil Defense	65	77.5%
	18	Saudi Press Agency (SPA)	114	67.5%
M-Business	**19**	Al-Rajhi Bank	8	91.3%
	20	Albilad Bank	42	81.3%
	21	Saudi Electricity Company (SEC)	22	85.6%
	22	Saudi Automotive Services Co. (SASCO)	20	87.5%
	23	E-Mall	142	40.0%
	24	eXtra Stores	1	96.3%
	25	Othaim Markets	40	81.6%
	26	Jarir Bookstore	138	52.5%
	27	Panda	59	78.4%
	28	Saudi Company for Hardware (SACO)	48	80.3%
	29	Mobily for Individuals	52	80.0%
	30	Saudi Telecom Company (STC)	2	96.3%
	31	al-jazirah newspaper	67	77.2%
	32	Al-Riyadh newspaper	24	85.3%
	33	Arab News	66	77.5%
	34	Aleqtesadiah newspaper	14	89.4%
	35	Al Tayyar Travel	135	58.1%
	36	Solidarity Saudi Takaful	88	74.1%

Difference between the sample and the population were examined using a t-test. The t-test results showed that there were no significant differences between the sample and the population (t_{1134}= 1.273, cv=1.960, p>0.05).

The sample was utilized to run several statistical tests for the different evaluation factors, in order to formulate hypotheses. The first t-test was used to examine the difference between M-Business and M-Government. The t-test results showed that M-Business adoption was of higher quality than that for M-Government (t_{1134}= 3.813, cv=1.960, p<0.001). The t-test results showed that M-Business was of higher quality than M-Government in terms of usability (t_{1134}= 2.679, cv=1.960, p<0.001) and attractiveness (t_{1134}= 3.422, cv=1.960, p<0.001). The t-test results also showed that the content in M-Business was richer (t_{1134}= 3.856, cv=1.960, p<0.001) and more organized (t_{1134}= 3.599, cv=1.960, p<0.001) than that in M-Government. Therefore, three hypotheses of mobile computing in Saudi Arabia were proposed:

H1: M-Business is more progressive in mobile computing than M-Government.

H2: M-Business is more usable and attractive than M-Government.

H3: M-Business content is richer and more organized than M-Government.

This study has identified several implications that are of relevance to the theory and practice of mobile computing. For researchers, it provides insight into the similarities and differences between M-Government and M-Business practices in Saudi Arabia. It supports further studies in the mobile computing field by proposing a sample of mobile applications that represents the population with statistical significance. For practitioners, the study showed trends and general characteristics of mobile computing practice in Saudi Arabia. For example, it was noteworthy that displaying content was the most dominant purpose of mobile computing adoption, yet the evaluation of the content displayed was rated lower than other factors. It also revealed dominant mobile platforms, by explaining major mobile presence purposes and showing key evaluation factors, such as usability, organization, content and visual appearance. The experience gained from this study suggested that Saudi organizations should embrace the development of mobile computing by considering the general trends and key characteristics of mobility practices. Further research in this area could be taken in different directions, such as usability heuristic evaluation of mobile computing and user

acceptance testing of mobile applications in both government and private businesses.

VI. Conclusion

Recent advances in the theory and practice of IT indicate that mobility is an emerging trend in the computing field. This study investigated the major trends in mobile computing by providing insight into key characteristics and quality levels of mobile computing practices in Saudi Arabia. Based on a four-phase process, mobile applications developed by Saudi organizations both governmental and in private business were evaluated by a focus group of eight, all of which were experts in the fields of mobile computing and usability engineering. The output from this study included a two-level categorization hierarchy of mobile applications in Saudi Arabia. Furthermore, the study proposed a representative sample of the population of mobile applications in Saudi Arabia which included equal numbers for M-Government and M-Business practices, with eighteen applications each. The study also formulated three hypotheses of the differences between M-Government and M-Business practices in Saudi Arabia, these hypotheses could be utilized to direct further research in the mobile computing field. All of the mobile applications which were qualified during this study were ranked in order to demonstrate the variations between the different sectors. Implications for research and practice have been highlighted along with directions for further research.

References

[1] I. Clarke, "Emerging value propositions for m-commerce," *Journal of Business Strategies*, vol. 18, pp. 133-148, 2001.

[2] CITC, "CITC Annual Report for information technology (IT)," [Online] http://www.citc.gov.sa/English/MediaCenter/Annualreport/Documents/PR_REP_008Eng.pdf, accessed on 10 January 2014, 2012.

[3] H. Verkasalo, C. Lopaz-Nicolas, F. J. Molina-Castillo, and H. Bouwman, "Analysis of users and non-users of smartphone applications," *Telematics and Informatics*, vol. 27, pp. 242-255, 2010.

[4] S. Po, S. Howard, F. Vetere, and M. B. Skov, "Heuristic evaluation and mobile usability: Bridging the realism gap," in *Mobile Human-Computer Interaction-Mobile HCI 2004*: Springer, pp. 49-60, 2004.

[5] J. Kjeldskov and J. Stage, "New techniques for usability evaluation of mobile systems," *International Journal of Human-Computer Studies*, vol. 60, pp. 599-620, 2004.

[6] D. Zhang and B. Adipat, "Challenges, methodologies, and issues in the usability testing of mobile applications," *International Journal of Human-Computer Interaction*, vol. 18, pp. 293-308, 2005.

[7] T. Kallio and A. Kaikkonen, "Usability testing of mobile applications: A comparison between laboratory and field testing," *Journal of Usability Studies*, vol. 1, pp. 23-28, 2005.

[8] D. Franke and C. Weise, "Providing a software quality framework for testing of mobile applications," in *IEEE Fourth International Conference on Software Testing, Verification and Validation (ICST)* Montreal, Canada: IEEE, pp. 431-434, 2011.

[9] M. B. Alotaibi, "Determinants of mobile service acceptance in Saudi Arabia: A revised UTAUT model," *International Journal of E-Services and Mobile Applications (IJESMA)*, vol. 5, pp. 43-61, 2013.

[10] E-Government, "Government Agencies Directory," National E-Government Portal, [Online] http://www.saudi.gov.sa/, accessed on 15 January 2014.

[11] Tadawal, "Directory of Saudi Companies," Saudi Stock Market Exchange, [Online] http://www.tadawul.com.sa/, accessed on 15 January 2014.

[12] Ministry of Commerce and Industry (MCI), "Index of Saudi Businesses," [Online] http://www.mci.gov.sa/, accessed on 15 January 2014.

VII. Appendix A

TABLE III. The evaluated mobile applications, ranked based on overall evaluation score

Category	#	Name	Sub-category	Rank	Score
M-Government	**1**	Ministry of Commerce and Industry	Ministries	55	79.1%
	2	Ministry of Transport (MOT)	Ministries	58	78.4%
	3	Ministry of Foreign Affairs	Ministries	30	83.1%
	4	Ministry of Health	Ministries	15	89.1%
	5	Ministry of Higher Education	Ministries	98	72.8%
	6	Ministry of Islamic Affairs, Endowments, Dawah and Guidance	Ministries	107	70.0%
	7	Ministry of Labor (MOL)	Ministries	91	73.4%
	8	Islamic University	Education	119	66.3%
	9	King Fahd University of Petroleum & Minerals (KFUPM)	Education	130	59.4%
	10	Princess Nora bint Abdul Rahman University	Education	109	69.4%
	11	Qassim University	Education	103	71.3%
	12	Taibah University	Education	110	69.4%
	13	Jubail Industrial College	Education	117	66.9%
	14	E-Learning Deanship at Imam University	Education	77	75.3%
	15	Library at KFUPM	Education	79	75.0%
	16	Female Teaching Center at KSU	Education	49	80.0%
	17	Saudi Food & Drug Authority (Saudi FDA)	Authorities	111	68.4%
	18	Saudi Commission for Tourism and Antiquities (SCTA)	Authorities	122	63.1%
	19	Saudi Arabian General Investment Authority (SAGIA)	Authorities	125	62.2%
	20	Saudi Professional League (SPL)	Authorities	50	80.0%
	23	King Fahad Specialist Hospital (KFSH)	Health	123	62.5%
	24	King Fahad Medical City (KFMC)	Health	47	80.3%
	25	Saudi Physical Therapy Association (SPTA)	Health	81	75.0%
	26	Emergency Call Center by MOH (call 937)	Health	41	81.3%
	27	Technical and Vocational Training Corporation (TVTC)	Government corporation	129	61.3%

	28	Saudi Post Corporation	Government corporation	121	63.4%
	29	Saline Water Conversion Corporation (SWWCC)	Government corporation	23	85.3%
	21	Holy Makkah Municipality	Other agencies	72	75.6%
	22	Municipality of Eastern Region	Other agencies	80	75.0%
	30	National Information Center (NIC) at MOI	Other agencies	78	75.3%
	31	Riyadh Chamber of Commerce and Industry	Other agencies	69	76.6%
	32	Saudi Arabian Airlines (Saudia)	Other agencies	92	73.4%
	33	Saudi Arabian Cultural Mission in New Zealand	Other agencies	54	79.4%
	34	Saudi E-Government (Yesser)	Other agencies	12	90.0%
	35	General Directorate of Civil Defense	Other agencies	65	77.5%
	36	Saudi TV Channel 1	Other agencies	140	43.8%
	37	Quran Radio	Other agencies	33	82.5%
	38	Real-Estate Development Found at Eastern Province	Other agencies	25	85.0%
	39	Saudi Press Agency (SPA)	Other agencies	114	67.5%
	40	Critical Locations Service by Riyadh municipality (Hather)	Government services	108	70.0%
	41	Staff Services by KSU	Government services	93	73.1%
	42	Student Services by KSU	Government services	126	61.9%
	43	Postal Address Locator by Saudi Post	Government services	115	67.5%
	44	Tourism Navigator by SCTA	Government services	34	82.5%
	45	Geographic Notifications by Jeddah Municipality	Government services	94	73.1%
	46	Riyadh Municipality Notifications	Government services	73	75.6%
	47	Government Notifications (Ishaar) by MOI	Government services	86	74.4%
	48	Water Notifications by Ministry of Water & Electricity	Government services	112	68.1%
	49	Commercial Violation Notification Service by MCI	Government services	26	85.0%
	50	Riyadh Map (Wsef) by Riyadh Municipality	Government services	70	76.3%
	51	Calendar of Events by SCTA	Government services	104	71.3%
	52	Electronic Directory by MOH	Government content	100	71.9%
	53	Saudi School Curricula by MOE	Government content	68	76.9%
	54	National Plan for CIT (Tahawul) by MCIT	Government content	82	75.0%
	55	Saudi Shelter Guide by SCTA	Government content	71	76.3%
	56	Saudi Laws by Bureau of Investigation and Prosecution (BIP)	Government content	101	71.9%
	57	Riyadh Directory by Arriyadh Development Commission	Government content	83	75.0%
	58	Jeddah Explorer to Licenses of Digs	Government content	120	63.8%
	59	Riyadh Book Fair by Ministry of Culture and Information	Government content	95	73.1%
	60	Index of Consumer Goods by MCI	Government content	99	72.5%
	61	Indicators of Higher Education by MOHE	Government content	74	75.6%
M-Business	**62**	Al-Rajhi Bank	Banks & Financial Services	8	91.3%
	63	National Commercial Bank (NCB)	Banks & Financial Services	21	85.6%
	64	Albilad Bank	Banks & Financial Services	42	81.3%
	65	Alinma Bank	Banks & Financial Services	6	92.2%
	66	Arab National Bank (ANB)	Banks & Financial Services	3	93.4%
	67	Banque Saudi Fransi	Banks & Financial Services	45	80.6%
	68	Riyad Bank	Banks & Financial Services	16	88.8%
	69	Saudi British Bank (SABB)	Banks & Financial Services	56	78.8%
	70	Saudi Investment Bank (SAIB)	Banks & Financial Services	51	80.0%
	71	Samba	Banks & Financial Services	4	92.5%
	72	Saudi Hollandi Bank	Banks & Financial Services	10	90.6%
	73	Aljazera Bank	Banks & Financial Services	9	91.3%
	74	Saudi Fransi Capital	Banks & Financial Services	17	88.8%
	75	Al Rajhi Capital	Banks & Financial Services	11	90.6%
	76	Mobily Mubasher	Banks & Financial Services	27	85.0%
	77	Samba Tadawul	Banks & Financial Services	28	85.0%
	78	Saudi Hollandi Capital	Banks & Financial Services	31	83.1%
	79	Tadawul	Banks & Financial Services	18	88.1%
	80	Saudi Electricity Company (SEC)	Energy & Utilities	22	85.6%
	81	Saudi Automotive Services Co. (SASCO)	Energy & Utilities	20	87.5%
	82	Saudi Lighting Company	Energy & Utilities	131	59.4%
	83	E-Mall	Retail & Commerce	142	40.0%
	84	eXtra Stores	Retail & Commerce	1	96.3%
	85	Othaim Markets	Retail & Commerce	40	81.6%
	86	Jarir Bookstore	Retail & Commerce	138	52.5%
	87	Obeikan Store	Retail & Commerce	96	73.1%
	88	Panda	Retail & Commerce	59	78.4%
	89	Saudi Company for Hardware (SACO)	Retail & Commerce	48	80.3%
	90	Bin Dawood Markets	Retail & Commerce	134	58.1%
	91	Mobily for Business	ICT	87	74.4%
	92	Mobily for Individuals	ICT	52	80.0%
	93	Saudi Telecom Company (STC)	ICT	2	96.3%

	94	Zain	ICT	13	89.4%
	95	Panorama FM	Media and Publishing	46	80.6%
	96	Al Arabiya	Media and Publishing	5	92.5%
	97	MBC FM	Media and Publishing	60	78.1%
	98	MBC NOW	Media and Publishing	75	75.6%
	99	al hayat Newspaper	Media and Publishing	37	81.9%
	100	Alyaum Newspaper	Media and Publishing	43	81.3%
	101	al-jazirah Newspaper	Media and Publishing	67	77.2%
	102	Al-Riyadh Newspaper	Media and Publishing	24	85.3%
	103	Arab News	Media and Publishing	66	77.5%
	104	Arriyadiyah	Media and Publishing	89	73.8%
	105	Sabq	Media and Publishing	29	83.8%
	106	Saudi Gazette	Media and Publishing	61	78.1%
	107	Aleqtesadiah Newspaper	Media and Publishing	14	89.4%
	108	Hasa Newspaper	Media and Publishing	139	47.5%
	109	Arriyadiyah Newspaper	Media and Publishing	57	78.8%
	110	Alweeam Newspaper	Media and Publishing	84	75.0%
	111	Al Tayyar Travel	Other businesses	135	58.1%
	120	Mahd alburaq	Other businesses	141	43.1%
	112	Aljomaih Automotive	Other businesses	132	58.8%
	113	Solidarity Saudi Takaful	Other businesses	88	74.1%
	114	Tawuniya	Other businesses	62	78.1%
	115	Munch Bakery	Other businesses	124	62.5%
	116	Alromansiah Restaurants	Other businesses	32	83.1%
	117	Effat University	Other businesses	127	61.9%
	118	SABIC	Other businesses	116	67.5%
	119	Spring Rose	Other businesses	118	66.9%
	121	Al-dawaa Pharmacy	Other businesses	53	80.0%
	122	Islamic Development Bank Group Business Forum - THIQAH	Business content	128	61.9%
	123	Abyat by STC	Business content	97	73.1%
	124	Argaam	Business content	35	82.5%
	125	Asrar STC	Business content	90	73.8%
	126	Government Directory by Mobily	Business content	105	71.3%
	127	Deeni Yaqini by STC	Business content	44	81.3%
	128	Hajj Application by Mobily	Business content	63	78.1%
	129	Donation by STC	Business content	113	68.1%
	130	Saudi Green Building Forum	Business content	133	58.8%
	131	Quality Forum by SEC	Business content	136	56.3%
	132	Your Health	Business content	106	71.3%
	133	Your Eye Health	Business content	137	54.4%
	134	Iktissab by Othaim Markets	Business content	76	75.6%
	135	Lamsa	Business content	36	82.5%
	136	STC Library	Business content	19	88.1%
	137	Alhilal Football Club by Mobily	Business content	64	78.1%
	138	Dawri Plus by STC	Business content	7	91.9%
	139	Aramco Job	Business services	102	71.9%
	140	SABIC Materials Selector	Business services	38	81.9%
	141	SABIC PP Finder	Business services	39	81.9%
	142	Rassid by SEC	Business services	85	75.0%

Digital Divide or Digital Disinterest?

By Dr. Elliott S. Lynn
American Public University

The digital divide among minorities remains a hot topic within circles in academia and often correlated with lack of resources and opportunity; however, there are many with resources and opportunity that have no desire or intention to become more computer literate than social networking sites require. Is there a divide created within minority communities that have nothing to do with lack or substandard access to resources? In many cases, yes,

I conducted a study on the correlation between computer literacy levels among African American men and the relationship to pursue Information Technology related careers and the results were startling. My original expectation was to find low computer literacy levels because of an obvious lack of resources and lower income levels, but to my surprise there was a greater correlation between low computer literacy levels and a lack of interest of in information technology, negative views toward information technology professionals, and no sense of urgency of becoming computer literate.

One of the contradictory findings to previous digital divide studies was that 98% of the African American male population surveyed had at least 4 hours of computer access per day. In fact,89.6% of participants owned computers with almost half of those owners having more than one computer in the household. The evidence suggested there is no lack of technology related resources among this population. These findings in this study alone do not indicate there is not a digital divide, but other issues besides resources creating a computer literacy gap among minorities.

In the study, many of the participants were computer literate, educated, and had access to a computer, but only 34% of participants had positions that required some knowledge of a PC. Only 8% of the participants worked in IT based positions. While 16% of participants indicated that their computer abilities influenced their career path, over 80% of participants indicated their computer abilities had no influence on their career path. The evidence suggested that African American males are not opting for IT related careers because of reasons beyond computer literacy and lack of resources.

Researchers suggest there is evidence that African American males significantly lack computer literacy skills when compared to other nationalities because of limited computer and Internet availability, also known as the *digital divide* (Malveaux, 1999). Valadez and Duran (2007) explained the digital divide based studies show White and Asian Americans have more than 20% greater access to computer and internet usage causing societal consequences. Carver (1994) acknowledged the lack of computer literacy skills and computing resources among African American men will hinder them from being a part of the information age and making any significant contribution to information-based societies. Computer literacy levels among African American have not increased in the proportion of internet and computer usage. Failure to increase computer literacy levels among African American men in an innovative society forces them out of

the job market because of an inability to meet computer literacy demands of the workplace. Cross (2003) asserted computer literacy for African Americans is an absolute necessity in order to compete successfully in the work force of the 21 century. It is important to note that African Americans are not the only ethnicity that needs computer literacy skills in the workplace, but research on the computer literacy levels shows statically significant differences among African American men. Hawkins and Paris (2007) explained the documented research of the digital divide is a strong indication of African American underachievement in the 21st century technology based workplace. African American men are at a disadvantage compared to other social groups seeking employment because of their lack of resources and computer literacy training. According to Smith (2005), "White households (46.1%) are twice as likely to have internet access in comparison to African American (23.5%) households" (p. 14). African American men do not have as much access to utilize computers as much as White and Asian groups. A lack of confidence with information technology may hamper academic and career success (Smith, 2002, p. 1). A lack of confidence in technology is one of several reasons cited for the small number of African American men in technology related professions (Haynes 2000; Smith, 2002). Some studies suggest this lack of confidence with technology among African American men is a result of the digital divide in the United States.

As the need for technology evolved from a luxury to a necessity in 21st century society studies began to emerge concluding disproportionate differences between White Americans and African Americans computer usage and skills (Hacker and Mason, 2003; Jackson, Von Eye, & Wenger, 2002; Koss, 2001). Malveaux (1999) defined the digital divide as the difference between African Americans greater lack of access to the internet and computing technology compared to other races. Hacker and Mason (2003) acknowledged that African Americans were 18% below the national average for home internet access. The lack of resources and access to the internet and computing technology among African American men affects their ability to make a difference in a competitive technological workplace. Loube (2003) stated, "Internet usage and computer ownership is already becoming essential for participating in the information economy" (p. 437). Without the fundamental understanding of how the internet or computer technology works, there is a distinct disadvantage to African American men with a desire to be a productive element in a technologically dependent society. For example, several organizations have abandoned traditional applications and replaced them with a website or a computer terminal that requires some level of computer knowledge to complete the application. Kuhn and Skuterud (2004) explained less access to computer and internet resources correlate with longer unemployment rates and durations among African Americas. Many African Americans do not have access to a computer or technology based resources outside of public schools creating a larger problem for some African American adults that have graduated or no longer attend public school. Brown (2000) explained that whites have more access and greater use of technology than African Americans increasing their likelihood of failing to reach their full potential in a technology dependent society. Koss (2001) asserted that in the 21st century, the United States has more computers than any country in the world but the lack of computer and internet availability among African Americans continues to rise affecting economic growth opportunity. Low levels of education among African American men correlates

with the inability to gain knowledge of computers and computing skills, especially if public schools are the only access to computers they have.

Tumposky (2001) argued some African Americans have equal access to computers and technology but there is a disinterest in technology because of a lack of positive messages about technology in their pop culture. There are many movies, images, and negativisms associated with the use of technology that discourages some African American males from using technology. For example, the role of someone using a computer in television programs or movies associate that person as being a *nerd* or someone that some African Americans youth may not view as "cool". Kendall (1999) concluded that some African American men had a negative view of computer skills at times associating computer skills with masculinity and homosexuality. Negative views and experiences with technology can push some African Americans to use technology only when they deem necessary or if there is a specific reason that would yield a benefit from its use. Jackson, Eye, Biocca, Barbatsis, et al. (2005) acknowledged in a sample of African American males with access to the internet and aspirations to pursue a career in entertainment or sports had less interest in utilizing the internet for academic purposes or proficiency because of perceived uselessness in obtaining their goal. This failure to recognize that technology is an essential tool for communication, networking, and increasing knowledge about their interest creates a self-imposed disadvantage, even when pursuing careers of a nontechnical nature.

I am not stating that the digital divide and or differences in low computer literacy levels among minorities are self-imposed or solely by choice, but a multi faceted problem beyond lack of resources and ability. This is a serious problem that will have long lasting effects on minorities in education, the workplace, and society as a whole.

SESSION

MANAGEMENT, TOOLS, WEB, VIRTUAL COMMUNITIES + APPLICATIONS, STANDARDS, POLICIES

Chair(s)

TBA

Novel Design Approach to Build Audit Rule Ontology for Healthcare Decision Support Systems

Numanul Subhani, Robert Kent
School of Computer Science
University of Windsor, Windsor, Ontario, Canada

Abstract— *Continuous Auditing (CA) has been investigated over time and it is, somewhat, in practice within financial and transactional auditing as a part of continuous assurance and monitoring. Enterprise Information Systems that run their activities in the form of processes require continuous auditing of a process that invokes the action(s) specified in the policies and rules in a continuous manner. A shared vocabulary, or common ontology, defines the processes and the audit rule ontology for processes or modules are integrated to form a hybrid ontology that supports the acquisition and evolution of ontologies. In this paper, we propose a methodology to construct a Healthcare Common Ontology by coupling to an expert system. We also propose a direction to generate an audit rule ontology of a process or module from defined rules for the purpose of Continuous Process Auditing (CPA) for Healthcare Decision Support Systems.*

Keywords: Semantic Web, OWL, Audit Rule Ontology, Healthcare Decision Support Systems, Enterprise Information Systems

1. Introduction

Using straightforward terms, auditing encompasses a variety of methods used to measure the compliance of a system to defined rules and guidelines. Auditing is just one facet of a more extensive set of process, often based on accounting, but extending to support assurance that the system is working as intended. Clearly, auditing is applied in many domains, from government to business, education and health care, among many others. Underscoring the breadth of auditing applications is the fact that most auditing is still performed by human agents, professionally trained and experienced in many aspects of evidence gathering, interpretation of rules and guidelines, and clarifying of final reports in respect of limitations.

Increasingly, complex systems have grown beyond the capacities of human driven auditing to perform meaningful audits in a timely fashion that serves stakeholders and oversight bodies. Such systems encompass networks of human agents and also highly automated software systems with semi-autonomous sub-systems, all of which are assumed to be imperfect or vulnerable to risk. To this end, researchers have focused attention on the notion of automating significant parts of auditing, both as embedded components within systems working autonomously, and as decision support components serving human analysts.

One vital element throughout auditing concerns knowledge, namely, its acquisition, interpretation and uncertainty, or vagueness, in reasoning. From the outset, auditing practice dictates that meaningful definitions must be determined and documented, for the system to be audited, components and processes within the system, evaluation measures, rules and actions to be applied, and limitations or constraints. In all aspects, auditors must work with suitable knowledge expressed in natural language terms for human consumption, but also expressed in terms appropriate for application and reasoning through logic, using computational techniques in particular.

Knowledge is an essential part of most Semantic Web applications and ontology, which is a formal explicit description of concepts or classes in a domain of discourse [10], is an essential part of the knowledge. Extracting knowledge from text in a semi-automatic way and identifying effective procedures for achieving useful and reliable results are challenging and daunting scientific research areas. In the auditing field, most audit rules are defined in the context of human understanding and language and can be used to support human cognitive reasoning. Inference based on interpretations of those audit rules are as essential to the Semantic Web as application domain ontologies. Ontology-based reasoning has known shortcomings and limitations compared with rule-based reasoning [12]. To represent inferential knowledge, ontology alone is insufficient [11], but inferential rules are an essential part of the knowledge in an audit rule ontology for a process or module in Continuous Process Auditing (CPA) for real-time Decision Support systems [4].

Berners-Lee [5] defined the semantic web as an extension of the current content-based web, in which information is given well-defined meaning. An ontology is usually defined as a formal specification of domain knowledge conceptualization. Ontology learning by application of semi-automatic methods has been studied and most techniques are from free natural language text. Though chronological, topological, and other types of semantic relation already exist [25], in these methods only hierarchical concepts are extracted and reduced sets of semantic relations are in use.

In Continuous Process Auditing methodology, an audit rule sheet is defined for each process or module. The

matter of continuity of application in CPA ranges from continuous time-dependent modeling to discrete time steps of audit application adapted to application requirements. For approaches with small time steps, the sheer magnitude of computational tasks to support CPA demands use of coarse-grained analysis of sub-systems, and estimation techniques based on limited rule sets. This consideration is used to determine the degree of conceptualization as knowledge, audit measures using sensors and reasoning through rules and inference.

According to Wache [28], the hybrid approach to integrating ontologies is the most relevant approach because it allows for semantic heterogeneity and flexibility. Healthcare systems are complex and heterogeneous in nature and their data sources are semantically heterogeneous. A common ontology approach is straightforward for dealing with heterogeneous semantic data sources. Hybrid approaches and multiple ontologies to the heterogeneity problem of ontologies have been discussed [16]. Integration and making mappings of ontologies are necessary for this kind of scenario. In the context of this paper, a common ontology can be constructed from multiple Process Ontologies (PO) with integration of semantically heterogeneous data sources. A Mapping Ontology (MO) allows mappings between the various PO and it also seeks and relates mappings between the various schema of the integrated data. To represent the mappings automatically, or in a semi-automatic way, an expert system like JESS [8] is merged with the PO to infer new relations from the existing concepts.

The objective of this paper is to present a hybrid ontology construction approach that combines the common ontology (shared vocabulary) with multiple, integrated and mapped POs. This common ontology would be the top-most layer in topological semantic relations. Audit Rule Ontologies for Processes (AROPs) would be stemmed as the second layer, under the common ontology layer in CPA, to infer the audit rules. The rest of the paper is organized as follows. A general description and definition of CPA, Audit rules and domain ontologies in the context of healthcare decision support systems is provided in section 2. Section 3 illustrates our proposed approach. The semi-automated development of the Common ontology in the context of healthcare based on the proposed approach is presented in section 4. Section 5 discusses the possibilities of how an AROP can be generated from audit rules, then how to stem AROPs as a second layer under common ontology. Conclusion and future research directions are drawn in section 6.

2. Audit Rules and Ontologies

As defined by Gruber, an ontology is an explicit specification of conceptualization [10], that can serve as an effective and powerful tool to capture, store and work with domain knowledge in knowledge-based information systems. In terms of knowledge representation, there are several types of ontology, including high-level, generic, domain and application. Generally speaking, domain ontologies are intended to specify the conceptualization of particular real-world domains. Domain ontologies usually describe a set of concepts and activities related to domains such as finance, commerce or industries involved in the production or delivery of goods and services, and other examples.

Below we describe the ontology aspects relevant to processes, then audit rules and finally hybrid approaches.

2.1 Process Ontology (PO)

All activities in a process are linked as sequential steps either defined by higher business modelers or discovered by various established methodologies, such as workflow mining from labeled and unlabeled event logs, stochastic workflow analysis, rule-based approaches and so on. There are two system approaches to the study of any system and its behavior: the micro system (μ), studies the algorithms, sensors for collecting data, and atomic devices; and the macro system (M), which studies and models large systems composed of large numbers of algorithms and devices. Fig. 1 depicts the pictorial view of a process traversal in the macro system (M). For example, Process P1 traversed through $\mu_1, \mu_5, \mu_2, \mu_4$ respectively then terminated at μ_3 in a macro system (M), which can be viewed as directed acyclic graph (DAG). Discovering process in a macro system, a NP-hard problem, is similar to find a topological ordering of a process in a given DAG. Our focus is on already discovered or predefined; solving the problem of discovering processes in a macro system will be investigated in future research within the context of Ontology Evolution.

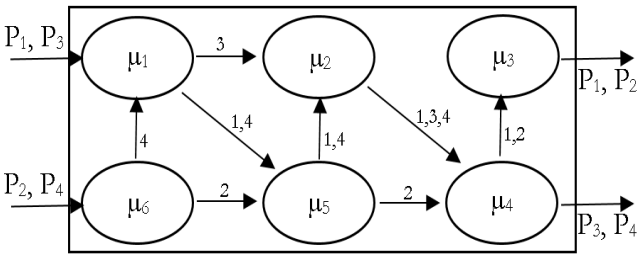


Fig. 1: How a Process is traversing in a Macro System (M). For example, Process P2 traversed through μ_6, μ_5, μ_4 respectively then terminated at μ_3.

Process Ontologies (PO) are constructed for each process with their defined concepts and databases that might be either homogeneous or heterogeneous in nature, and an expert system for PO mappings is coupled to construct our proposed hybrid layered ontology approach addressed in more detail in Section 3.

2.2 Audit Rules and Audit Rule Ontology of a Process (AROP)

Audit rules are defined for an activity or a system component within the audit scope in any human-performed auditing mechanism, as a first step of common auditing procedures. Audit rules are based on both the hierarchical structure of organization and the enforced business controls. The same audit rules discernment and definition approach can be applied and implemented in any Continuous Process Auditing system where a process has to traverse through various components by applying audit rules sequentially.

An Audit Rule Ontology for a Process (AROP) would be used to detect exceptions to the audit rules in a process during CPA. Semantic rule-based reasoning would facilitate construction of AROP in a semi-automatic way. AROPs would be used as second layer under common ontology in hybrid layered ontology model. We assume that human approval of all audit rules is enforced; the issue of autonomous automated approval through artificial intelligence is not considered in this discussion.

2.3 Ontology Design: A Hybrid Layered Model

Domain ontologies may be divided into linguistic and conceptual ontologies. According to Gruber [10], Conceptual Ontologies (CO) represent the domain objects, distinguishing between the primitive concepts and the defined concepts, whereas Linguistic Ontologies (LO) define words or contextual usages of words. The Process Ontology (PO) contains only the defined concepts and the Mapping Ontology (MO) contains both the defined and the underlying primitive concepts. The observation in [13] led to identifying some relationships between POs, MOs and LOs. Mappings between POs may be defined in terms of equivalence operators of some MO. The various meanings of words in MO references may be defined by LOs and this reference would provide a basis for formal, and exact or uncertain reasoning, and automatic translation of context-specific terms.

3. Proposed Hybrid Layered Approach

We propose a single common ontology approach with multiple POs for domain ontology construction. Each PO is attached to a database that might be heterogeneous in nature with other databases. Each PO describes the semantics of data sources individually. Inter-PO mapping is realized by the MO, which is defined with primitive concepts. The simplicity and flexibility permitting addition of new sources (like new POs) with little or no need of modification, is the main advantage of this mechanism. To integrate several POs addressing the same domain, this mechanism exploits the MO's capability to define equivalent and similar concepts. We discuss three mapping use cases of semantic integration:

a) Discovery of Mapping: To find the similarities and determine the concepts and properties for representing similar notions between two POs, we use PO structures, definitions of concepts, and instances of classes.

b) Mappings Representation of MO: To represent the mappings between two POs to enable reasoning with mappings. The mappings representation of inter-PO is used in defining the MO. The MO can consist of the OWL constructor or the equivalence relation defined, or we can use different inference engines to assert the MO automatically, starting from the logical rules between the concepts belonging to different PO.

c) Mapping Uses: How to define the mapping between PO is not described or not a goal in itself, but it can be done either automatically or interactively. The resulting mappings are used for various integration purposes, such ase answering the usersâĂŹ requests. We use JESS instructions language for this purpose.

As detailed by Wache [28], the first step is to develop the Healthcare Common Ontology using multiple POs in the top-most layer, then Audit Rule Ontologies of Processes would be stemmed as a second layer under the top-most layer to form a Hybrid Layered Ontology. Fig. 2 visualizes the conceptual model of both the layers alongside with the technologies would be used to integrate and to develop the whole operable Audit Rule Ontology system. More abstract description of construction, development and operational mechanism are discussed in the next section.

4. Development of Healthcare Common Ontology

Many business changes have occurred throughout history and a recent major change is related to Enterprise Systems (ES) and the methodology Enterprise Resource Planning (ERP) [14]. These approaches have transformed the way business data is collected, stored, disseminated and used [24]. Enterprise resource planning systems are defined as "information system packages that integrate information and information-based processes within and across functional areas in an organization" [17]. An ERP system is an enterprise system that affects many or all departments of a company. Though research of Continuous Auditing started in 1991 [7], [26], [27]; Kent et. al. [15] first envisioned the application of Continuous Auditing for Healthcare Decision Support Systems in 2010.

Ontologies have been used to represent knowledge and to help knowledge inference in clinical research [1], decision support and maintenance of system [22]. Jean et al [13] showed the specification of an ontology as a domain model allowing solutions for various issues in data indexing, data exchange and data integration. An ontological representation

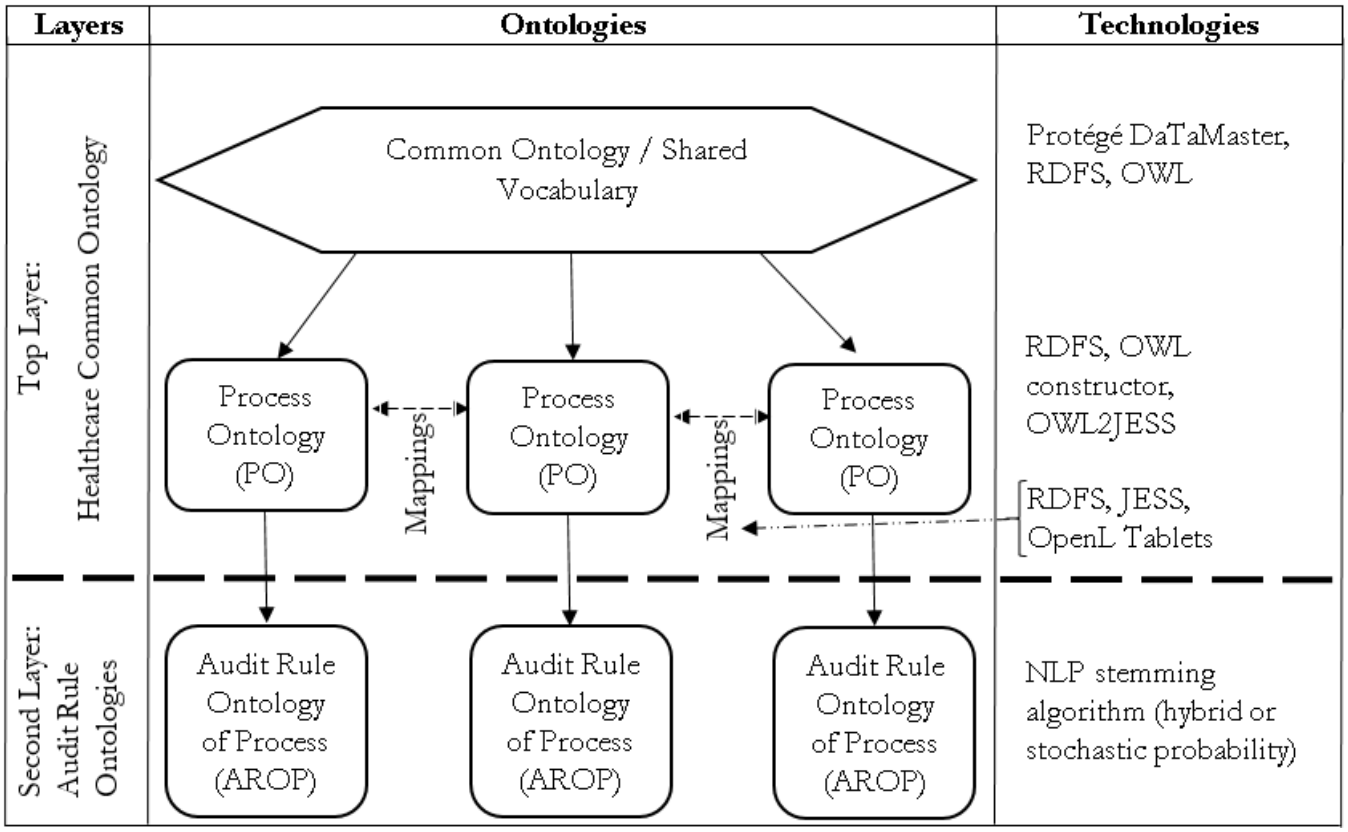


Fig. 2: Conceptual Model of Hybrid Audit Rule Ontology.

with rule-based reasoning as model for development of clinical decision support system was presented by Archour et al [1]. Alles et al discuss Continuous *Process* Auditing (CPA) in enterprise system environments related to Healthcare Decision Support System, requires definition and discovery of processes [2], [3]. Each process can be constructed as a Process Ontology (PO), as described in the previous section.

Each data source (DS) containing the process knowledge or information is described by a PO. For example, PO-DS1 and PO-DS2 for the first and second data source, respectively, with the ability to obtain access to identifiable POs. All available data sources are merged to construct a common domain ontology. The common vocabulary of each data source becomes a sub-group of the common domain ontology.

The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. Shared data representations such as eXtensible Markup Language (XML) provides an elemental syntax for content structure within documents lies in the bottom of the Semantic Web Stack. Resource Description Framework (RDF) and RDF Schema (RDFS), a general method for describing information are on top of XML in the stack respectively. Web Ontology Language (OWL) [18], a family of knowledge representation languages lies on top of RDFS in the same semantic web stack, adds more vocabulary for describing properties and classes: among others, relations between classes (e.g. disjointness), cardinality (e.g. “exactly one"), equality, richer typing of properties, characteristics of properties (e.g. symmetry), and enumerated classes. OWL provides the tools for semantic reasoning to describe or to represent knowledge. The OWL constructors, and the equivalent relation (such as Rule Interchange Format - RIF) or the specific relation of the domain to be created between concepts belonging to different POs would be the MO operators.

4.1 Healthcare Common Ontology Conceptualization

The Protege Plug-in DaTaMaster [20] is used to implement the proposed solution. Protege Plug-in DaTaMaster imports data source schemes and their contents under OWL. It permits the integration of various data sources in single ontological representation. To implement the solution a) any data connection driver like ODBC or JDBC can be used to connect with data bases b) selection and visualization of table content by the user preference, and c) each activated and visualized table is transferred into a class or sub-class depending on the choice of user.

4.2 Healthcare Common Ontology Construction

The insertion of OWL constructors, relations as well as annotations participate in the process of semantically enhance data belong to different PO as well as to solves syntaxes and semantic heterogeneity of integrated systems, improving data exchange between them. We assigned a unique space name to each PO to resolve conflict context. A pre-tagging concept is used to pre-tag with the same Uniform Resource Identifier (URI) for all the classes, attributes and the instances belonging to one PO. We create relations to resolve the naming conflict. Two classes issued from two different data sources (different PO), where first one describes the designation or equipment (id, MRI machine, ER room, pharmacy, diagnostic result) and the second details a patient cases (id, symptoms, current status, drugs). Both classes treat the same patient or equipment but the semantic of their data sources are different. A manual relation was created between the both classes that maintains an equivalent relation for both instances. For large ontology, automatic definition of the MO can be implemented with JESS [8] instructions to discover the common attribute.

4.3 Healthcare Common Ontology Operation

A rule based reasoning engine like JESS [8] can be used with the ontology instances. Tools like OWL2JESS [19] and OpenL Tablets [21] facilitate the necessary conversion of the OWL ontology code to facts and rules. RDFS and OWL verified coherence and uniformity of the ontology will permit us to design, evaluate and refine the original obtained ontology. Three layers of knowledge: the ontology model layer, the ontology layer and instances layer encapsulated by the obtained knowledge base.

OpenL Tablets is an open source rules engine and rules repository tool [21] which has Java Wrappers that can integrate and enable interoperability with JESS rules and instructions. The exploitation of the ontology is ensured using rules and requests permitting fetching from the knowledge base through a set of JESS commands, such as “defrule" and “defquery"

4.4 Towards Audit Rule to AROP Generation and Stemming of AROPs

So far we have designed and implemented the Healthcare Common Ontology as the top-most layer in a Hybrid Layered Ontology towards the generation of Audit Rule Ontology of a Process as a second layer. We have planned to devise a specific semantic similarity technique and we have been mechanizing a natural language processing (NLP) technique, like stemming algorithms, to add the AROPs to the Healthcare Common Ontology. Stemming, also known as branching, is the process for reducing inflected or derived words to their stem or root âĂŞ generally a written word form. Stemming is also a very powerful technique for mapping related words to the same stem or root. In particular we are considering the hybrid approach or stochastic probability type of stemming algorithm to identify the root or stem from the top-most layer Healthcare Common Ontology to generate the AROPs and add the AROPs as a second layer.

5. Ontology Evolution

Ontology evolution is defined as the process of updating the previous ontology version, in order to take into consideration changes within the domain, reflected in its conceptualization or application [6]. The evolutionary process, a crucial part of the ontology lifecycle, creates newer versions with added stemming down the tree from the original ontology. Since the uniformity and coherence of the ontology must be respected, the evolution process is difficult to implement semi-automatically and should be considered as beyond human capacities for complex ontologies

5.1 Evolution Methodology

Changes Identification of the domain is the fundamental step to form an ontology evolution strategy. There are two main methods to identify the changes - descending and ascending. The definition of domain update or the update of the ontology usage methods would be the *descending identification*. The changes that identified from the ontology analysis itself, i.e. by using heuristic rules or statistical inference for optimization, are the *ascending identification*.

Editing ontology changes might be either elementary, intermediate or complex. Stanjanovic discussed the *elementary changes* as non-decomposable changes given by a suppression or adding of ontological entities [23]. *Complex changes* are composed of two or more elementary changes forms that together form a logical entity. Giorgos et. al. [9] described the complex changes may be composed of other types of changes that happen in between elementary and complex, and which may be called *intermediate changes*. The insertion of links belongings to the concept along with other existing concepts must be done after the addition of a new concept. The isolation of one or more concepts may be the reason for the concept suppression. All conceptual or semantic relations, linking the suppressed concept with other concepts must then be deleted along with linked instances. We are devising an algorithm to add the complex changes by focusing on the following steps: the checking of changes made, validation of changes, changes implementation, analyzing the compatibility between previous and newer version, and finally validating the newer version in the community.

5.2 Healthcare Common Ontology Evolution

The evolution of the addition, suppression and modification operations on the knowledge base using the JESS language “assert" for addition, “modify" for modification,

"retract" for suppression. It is compulsory that the development of a system which propagates the changes automatically. In order to make changes transparent, the system must guide the user during operation. The change may be implemented either on the ontological level or on the instances level. Adding, suppressing and changing a concept, a relation or a semantic relation and updating conceptual relations may implemented on the ontological level whereas adding, suppressing and changing an instance should be implemented on the instances level.

An ontology evolution repository may be maintained for previous version and the process the changes made to the ontology. The repository would preserve the history of different versions of the ontology which helps to make changes forward and backward compatible.

6. Conclusion

A domain ontology such as Hybrid Layered Ontology is an approach that permits very efficient knowledge management and gives a unified conceptualization of the domain. In this paper an approach using multiple Process Ontologies based on knowledge extraction and integration from multiple, different data sources, which enhances semantically the final result, was presented. The semi-automatic ontology construction not only allows a faster and more efficient construction, but also may aid significantly in saving manual time consumption, effort and consistency. In order to define the Mapping Ontology automatically from the logical rules, an expert system JESS is integrated with OpenL Tablets. To update the knowledge base for better diagnoses and maintenance, a strategy for ontology evolution at conceptual, relational and instance levels are presented.

Acknowledgment

We acknowledge support from the Canadian Institutes for Health Research (CIHR).

References

[1] S. Achour, M. Dojat, C. Rieux, P. Bierling, and E. Lepage, "A umls-based knowledge acquisition tool for rule-based clinical decision support system development," *Journal of the American Medical Informatics Association*, vol. 8, no. 4, pp. 351–360, 2001.

[2] M. Alles, G. Brennan, A. Kogan, and M. Vasarhelyi, "Continuous monitoring of business process controls: A pilot implementation of a continuous auditing system at siemens," *International Journal of Accounting Information Systems*, vol. 7, no. 2, pp. 137–161, 2006.

[3] M. Alles, A. Kogan, M. Vasarhelyi, and J. Wu, "Continuity equations: Analytical monitoring of business processes in continuous auditing," *12th World Continuous Auditing Symposium*, 2006.

[4] R. Baksa and M. Turoff, "Continuous auditing as a foundation for real time decision support - implementation challenges and successes," *Supporting Real Time Decision-Making, Annals of Information Systems*, vol. 13, pp. 237–252, 2011.

[5] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific American*, vol. May, 2001.

[6] N. Casellas, *Methodologies, Tools and Languages for Ontology Design*. Springer Science+Business Media, Law, Governance and Technology Series 3,, ch. 3.

[7] D. Chan and M. Vasarhelyi, "Innovation and practice of continuous auditing," *International Journal of Accounting Information Systems*, vol. 12, pp. 152–160, 2011.

[8] E. Friedman-Hill, *JESS in action : Rule-Based system in Java*, isbn:1930110898 ed. Manning Publications, 2003.

[9] F. Giorgos, P. Dimitris, and A. Grigoris, "Evolving ontology evolution," *Current Trends in Theory and Practice of Computer science*, vol. SOFSEM 06, 2006.

[10] T. Gruber, "A translation approach to portable ontology specifications," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, 1993.

[11] F. Harmelen and D. Fensel, "Practical knowledge representation for the web," *Proc. 16th IntŠl Joint Conf. Artificial Intelligence*, 1999.

[12] I. Horrocks, "Daml+oil: A description logic for the semantic web," *IEEE Data Enggineering*, vol. 25, no. 1, pp. 4–9, 2002.

[13] S. Jean, G. Pierra, and Y. Ait Ameur, "A domain ontologies : A database-oriented analysis," *LNBIP: Web Information Systems and Technologies*, pp. 238–254, 2007.

[14] A. Kanellou and C. Spathis, "Auditing in enterprise system environment: a synthesis," *Journal of Enterprise Information Management*, vol. 24, no. 6, pp. 494–519, 2011.

[15] R. Kent, A. Zahid, and A. Snowdon, "Continuous auditing for health care decision support systems," *Intelligent Decision Technologies*, vol. SIST, no. 10, pp. 731–741, 2011.

[16] M. Klein, "Combining and relating ontologies: an analysis of problems and solutions," *Proceedings - IJCAI Workshop*, pp. 53–62, 2001.

[17] K. Kumar and J. Hillegersberg, "Enterprise resource planning experiences and evolution," *Common ACM*, vol. 43, no. 3, pp. 22–26, 2000.

[18] D. McGuinness and F. Harmelen. (2004-2009) Owl web ontology language overview. [Online]. Available: http://www.w3.org/TR/owl-features/

[19] J. Mei, E. Bontas, and Z. Lin, "Owl2jess: A transformational implementation of the owl semantics," *Lecture Notes in Computer Science*, vol. 3759, pp. 599–608, 2005.

[20] C. Nyulas, M. O'Connor, and T. Samson, "Datamaster - a plug-in for importing schemas and data from relational databases into protégé stanford medical informatics," *Stanford University School of Medicine, Stanford, CA 94305*, 2007.

[21] A. Ostrovski and et. al. (2004-2014) Openl tablets - easy business rules. [Online]. Available: http://openl-tablets.sourceforge.net/

[22] I. Rasovska, B. Chebel-Morello, and N. Zerhouni, "Process of s-maintenance: decision support system for maintenance intervention," *Emerging Technologies and Factory Automation*, pp. 8–15, 2005.

[23] L. Stanjanovic, "Methods and tools for ontology evolution," Ph.D. dissertation, 2004.

[24] S. Sutton, "Enterprise systems and the re-shaping of accounting systems: a call of research," *International Journal of Accounting Information Systems*, vol. 7, pp. 1–6, 2006.

[25] R. Valencia-Garcia and et. al., "An incremental approach for discovering medical knowledge from texts," *Expert Systems with Applications*, vol. 26, no. 3, pp. 291–299, 2004.

[26] M. Vasarhelyi and F. Halper, "The continuous audit of online systems," *Auditing: A Journal of Practice and Theory*, vol. 10, no. 1, pp. 110–125, 1991.

[27] M. Vasarhelyi, S. Romero, S. Kuenkaikaew, and J. Littley, "Adopting continuous audit - continuous monitoring in internal audit," *ISACA Journal*, vol. 3, no. 1, pp. 1–5, 2012.

[28] H. Wache and et. al., "Ontology-based integration of information - a survey of existing approaches," *Proceedings - IJCAI Workshop*, pp. 108–117, 2001.

New Windows Rootkit Technologies for Enhancing Digital Rights Management in Cloud Computing Environments

Woei-Jiunn Tsaur and Jia-Xin Wu
Department of Information Management, Da-Yeh University, Taiwan

Abstract - *With the increasing development of cloud computing technologies, digital contents such as music, movies, games, cartoon, and so on get more and more popular for entertainment, and therefore how to control and manage rights of digital contents has been becoming very critical in cloud computing environments. Sony-BMG used a rootkit to conceal the digital rights management (DRM) software, which is aptly installed in consumers' computers to prevent unauthorized copying. In order to prevent unauthorized users from deleting the DRM software by employing anti-rootkit tools to remove the rootkit, we develop a new Windows driver-hidden rootkit to enhance DRM in cloud computing environments. The proposed Windows driver-hidden rootkit has verified that it can successfully avoid detection of a variety of well-known anti-rootkit tools in cloud computing environments. Our contributions are mainly twofold. Firstly, in cloud computing environments we can use the proposed new rootkit technology to reinforce the Windows-based DRM software. Secondly, the stealth tricks of the proposed sophisticated Windows rootkit can be a great inspiration to defenders who need to effectively enhance the legitimate uses in cloud computing environments.*

Keywords: Digital Rights Management (DRM), cloud computing, rootkit, operating systems, system security.

1 Introduction

With the advent of new technologies such as cloud computing, it seems to cause problems of piracy even more than before. A huge amount of transactions of digital contents are expected under cloud computing environments. And the cloud computing environment is a place where many computers are available everywhere throughout the physical world connected seamlessly to the information systems. Therefore, it will be a more critical issue to control and manage rights of digital contents. Piracy and illegal distribution of digital contents are severe issues. Digital Rights Management (DRM) [1]-[7] aims at protecting digital contents from being abused through regulating the usage of digital contents. The DRM scheme is a digital protection technique that protects and manages the access rights of digital contents. It can prevent the confidential information of a digital content from unauthorized usages by illegal users.

Sony-BMG used a rootkit to conceal the DRM software, which is aptly installed in consumers' computers to prevent unauthorized copying. As stated in the literature [8], [9], rootkit is a stealth technology, and the intent with which this technology is used determines their malicious or otherwise legitimate purpose. The same technology used by rootkits is also used in security software such as firewalls and host-based intrusion prevention systems to extend the protection of the operating system. Therefore, the rootkit technologies may be employed in consumers' computers to conceal the DRM software for preventing unauthorized copying [7]. However, Tsaur's scheme [7] cannot be adopted in cloud environments because his proposed approach did not completely consider the properties of virtual machines and operating systems in cloud environments. In addition, cloud environments offer particularly attractive malware targets as they incorporate vast numbers of computing resources and high network bandwidths, and are increasingly becoming the operational home to many high-valued software systems and services. Attacks targeting clouds can provide significant chances to obtain control over resources and extract proprietary information. Thus, how to control and manage rights of digital contents has been becoming very critical in cloud computing environments.

In order to prevent unauthorized users from removing the rootkit of concealing the digital right management software by employing anti-rootkit tools in cloud computing environments, this paper is to propose new rootkit technologies for strengthening the DRM in protecting against the illegal distribution and consumption of copyrighted digital multimedia contents. In cloud computing environments, though many companies or organizations have been developing rootkit detectors to the public and undoubtedly they can detect known rootkits effectively, they cannot foresee what the result is when meeting unknown rootkits. In this paper, the proposed unknown Windows rootkit technologies are constructed in cloud operating systems, and have been verified that it can successfully evade detection and removal of a variety of well-known rootkit detectors. The contributions of this paper are mainly twofold. Firstly, we can use the proposed new rootkit technology to extend the reinforcement of the Windows-based DRM software in cloud computing environments, which can be a great inspiration to DRM software makers to effectively improve the current techniques of protecting against the illegal distribution and consumption of copyrighted digital multimedia contents. Secondly, the stealth tricks of the proposed subtle driver-hidden rootkit can be a great inspiration to defenders who need to effectively strengthen the legitimate uses in cloud computing environments.

The remainder of this paper is organized as follows. Section 2 surveys current rootkit creation and detection techniques.

Section 3 presents the method for developing new rootkit technologies in Windows cloud operating systems. Section 4 depicts the experimental results of testing the proposed rootkit's stealth ability in cloud operating systems. Finally, some concluding remarks are included in the last section.

2 Related work

There are essentially two different techniques that a rootkit can use to hide computer resources. One is hooking that intercepts the requests of accessing resources. The other is Direct Kernel Object Manipulation (DKOM) that manipulates the data used by operating systems to keep track of resources. The oldest kernel mode rootkit [8] uses table hooking to alter System Service Descriptor Table (SSDT) to hide processes, drivers, files, etc. Although it is a simple, stable and efficient method, it is easily detected by current rootkit detectors [10]. Hunt and Brubacher [11] introduced Detours, a library for intercepting arbitrary binary function. Later, this method is also applied by a rootkit, which replaces the first few instructions of a specific function with "jump" to point to the rootkit's code instead of targeting system tables. The aforementioned is named inline hooking. But VICE [8], a heuristic detector, is created to detect hookers no matter table hooking or inline hooking. In order to enhance inline hooking, brilliant rootkit makers combine a polymorphic technique [12] whose purpose is to generate different appearances of a piece of code. These appearances may look different but have the same functionality. On the other hand, Butler et al. [13] used DKOM to target EPROCESS, a kind of kernel data structures to record information related to a process, to alter an affiliated doubly linked list, and let the desired processes be hidden. When using DKOM, rootkit makers need to clearly understand the data structures in kernel, but it is more furtive than using hooking [8], [14]-[19]. The DKOM technique was first used in the FU rootkit and then used in FUTo to hide their drivers [8], [9]. In 2007, the DKOM-based Unreal rootkit was created and shown off that all of the famous detectors cannot detect it. However, at present several well-known detectors are capable of effectively detecting the above-mentioned three driver-hidden rootkits.

As for identifying rootkits, there are two main approaches to develop rootkit detectors. The first detection approach targets hiding mechanisms by detecting the presence of API hooking [14], [20]. It is similar to the signature-based detection [21], and thus it cannot catch unknown rootkits whose signatures of hiding mechanisms do not exist in its signatures repository in advance. The second approach targets hiding behaviors by detecting any discrepancies between the original and the fake. It collects resources information from two different storage places, and then compares each other to find rootkits. This approach is to belong to the cross-view rootkit detection [22]. It is noticed that in this approach both targeted information cannot be modified simultaneously by rootkits, otherwise a detector using this approach cannot distinguish the differences between the two places storing targets and then it must be useless. Although this method has the drawback, it does not need to maintain a signature database as used in the signature-based detection method.

3 Proposed new Windows rootkit technologies in cloud operating systems

In this section we propose new rootkit techniques which have abilities to escape well-known anti-rootkit tools in order to reinforce the Windows driver-format DRM in protecting against the illegal distribution and consumption of copyrighted digital multimedia contents. This paper discloses seven places, some of which may not be known by anti-rootkit developers, to hide driver information in cloud operating systems. The proposed new rootkit technologies are composed of seven tricks which will be presented in the following Items A-G, respectively.

A. Removing the Signature of PE (Portable Executable) Image

The PE file format is executable in operating systems environments and can be executed in multiple platforms. Almost all executable files (also including kernel mode drivers) are to use the PE file format which inherits the characteristics of the COFF (Common Object File Format) file format in UNIX operating systems. The PE file format contains five main parts: DOS MZ Header, DOS Stub, PE Header, Section Table and Sections, as shown in Fig. 1. Driver programs are loaded to memory via MmLoadSystemImage system function, and therefore their complete PE images exists in memory after loading. In order to avoid detectors to scan the PE files, the characteristics of the PE images must be removed. The method is to first remove the MZ Header, then the value of elfanew be set to zero in the DOS Header, and finally set the PE Header size to zero from pOptinalHeader and SizeofHeaders in the PE structure.

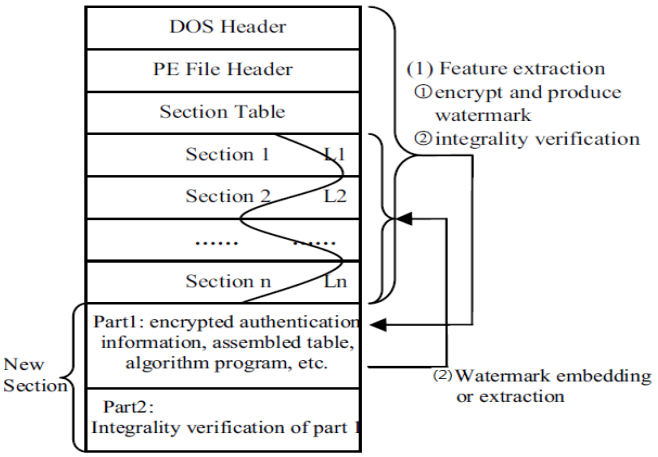


Fig. 1. The PE structure [23]

B. Removing Object Drivers and Object Devices from Object Directory

In the internal of Windows operating systems kernel, the

most fascinating part is objects. They contain all kinds of resources that are queried by kernel functions. The program of Windows object management is responsible to manage objects. All of them are kept in a tree of Object Directory whose definition is shown in Fig. 2. The Object Directory is established with several HashBuckets. Each one points to an Object Directory Entry structure whose definition is shown in Fig. 3. In Fig. 3, we can see that the Object member refers to an object, and ChainLink member points to another Object Directory Entry. Most of Object Drivers have at least one Object Device pointing to themselves, so both of them are needed to consider when a driver-hidden rootkit is invented. We can explore the whole Object Directory to find the desired Object Drivers and Object Devices, and then apply DKOM to achieve the purpose of hiding.

```
typedef struct Object_Directory
{
/0x000/   [37] Ptr32 _Object_Directory_Entry      HashBuckets
/0x094/   _EX_PUSH_LOCK                          Lock
/0x098/   Ptr32_DEVICE_MAP                       DeviceMap
/0x09c/   Uin4B                                  SessionId
/0x0a0/   Uin2B                                  Reserved
/0x0a2/   Uin2B                          SymbolicLinkUsageCount
}
Object_Directory  * pObject_Directory
```

Fig. 2. Definition of an Object Directory

```
typedef struct Object_Directory_Entry
{
/0x000/   Ptr32 _Object_Directory_Entry      ChainLink
/0x004/   Ptr32 Void                        Object
}
Object_Directory_Entry    * pObject_Directory_Entry
```

Fig. 3. Definition of an Object Directory Entry

C. Removing the driver-related information in the registry

After the rootkit has been loaded, in order to avoid that the system will be restored by restarting the computer, automatically loading rootkit on the system must be added to the SERVICE in the registry. It can be observed that after computer boot, the value of List existing in HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\ServiceGrouporder lists the order of loading drivers, as shown in Fig. 4. Rootkit detectors will be generally loaded after the value data "Boot Bus Extender", so the rootkit must be loaded before detectors so that it can successfully evade detection and removal of a variety of rootkit detectors. The rootkit then removes the driver-related information in the registry, such as Type, Start, ImagePath, ErrorControl, and DisplayName in Fig. 5, and thus rootkit detectors cannot scan for rootkit from the registry. In addition, the rootkit writes its value data to the registry when shutting so that it can be automatically loaded when a user restarts the computer.

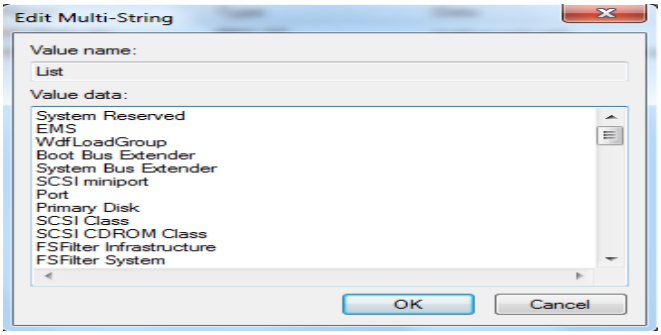


Fig. 4. Boot sequence in the registry

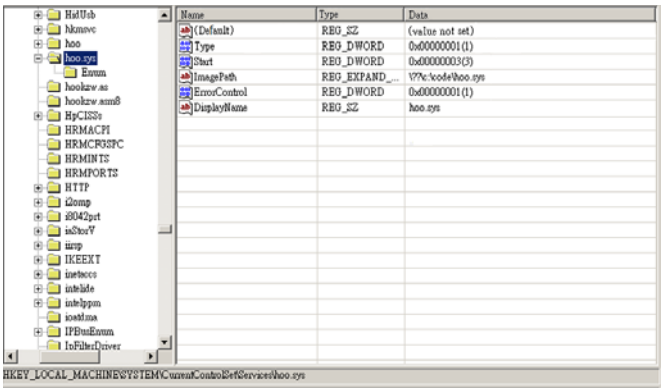


Fig. 5. Driver-related information in the registry

D. Removing Object Drivers from Driver's Object_Type

An Object_Type defines the common properties of the same kind of objects as shown in Fig. 6. Each kind of objects has a dedicated Object_Type. An Object_Type has a List_Entry data structure which keeps all of the same kind of Object_Creator_Info. The definition of Object_Creator_Info is described in Fig.7. Here the type of objects is referred to Object Driver whose definition is shown in Fig.8. We know that every object body is immediately preceded by an Object_Header structure whose Type member points to the Object_Type. We can exploit our loaded rootkit driver to get its Object Driver, then move to its Object_Header to get the pointer to the Object_Type, and finally check its TypeList member to find the desired Object Drivers to hide.

```
typedef struct Object_Type
{
/0x000/   _ERSOURCE                  Mutex
/0x038/   _List_Entry                TypeList
/0x040/   _UNICODE_STRING            Name
/0x048/   Ptr32 Void                 DefaultObject
/0x04c/   Uint4B                     Index
/0x050/   Uint4B                     TotalNumberOfObjects
/0x054/   Uint4B                     TotalNumberOfHandles
/0x058/   Uint4B                     HighWaterNumberOfObjects
/0x05c/   Uint4B                     HighWaterNumberOfHandles
/0x060/   _TYPE_INITIALIZER          TypeInfo
/0x0ac/   Uint4B                     Key
/0x0b0/   [4] _ERESOURCE             ObjectLocks
}
Object_Type     *pObject_Type
```

Fig. 6. Definition of an Object_Type

```
typedef  struct Object_Creator_Info
{
/0x000/   _List_Entry                  TypeList
/0x008/   Int4B                        CreateUniqueProcess
/0x00c/   Int4B                        CreateBackTraceIndex
/0x00e/   Ptr32 Void                   Reserved
}
Object_Creator_Info    * pObject_Creator_Info
```

Fig. 7. Definition of an Object_Creator_Info

```
typedef  struct Object_Driver {
/0x000/   Int2B                        Type
/0x002/   Int2B                        Size
/0x004/   Ptr32 Object_Device          DeviceObject
/0x008/   Uint4B                       Flags
/0x00c/   Ptr32 Void                   DriverStart
/0x010/   Uint4B                       DriverSize
/0x014/   Ptr32 Void                   DriverSection
/0x018/   Ptr32 _DRIVER_EXTENSION      DriverExtension
/0x01c/   _UNICODE_STRING              DriverName
/0x024/   Ptr32 _UNICODE_STRING        HardwareDatabase
/0x028/   Ptr32 _FAST_IO_DISPATCH      FastIoDispatch
/0x02c/   Ptr32                        DriverInit
/0x030/   Ptr32                        DriverStartIo
/0x034/   Ptr32                        DriverUnload
/0x038/   [28] Ptr32                   MajorFunction
}
Object_Driver      * pObject_Driver
```

Fig. 8. Definition of an Object Driver

E. Removing Object Devices from Device's Object Type

It is the same as the method described in Item D, except for different kind of objects. Here the type of objects is referred to Object Device whose definition is shown in Fig. 9. First, we find the Object Device through our loaded rootkit Driver. After getting the Object Device, we can find an Object_Type. Finally, we traverse its List_Entry to locate the desired Object Devices to hide.

```
typedef  struct Object_Device {
/0x000/   Int2B                        Type
/0x002/   Size                         Uint2B
/0x004/   Int4B                        ReferecnceCount
/0x008/   Ptr32 Object_Device          DriverObject
/0x00c/   Ptr32 Object_Device          NextDevice
/0x010/   Ptr32 Object_Device          AttachedDevice
/0x014/   Ptr32 _IRP                   CurrentIrp
/0x018/   Ptr32 _IO_TIMER              Timer
/0x01c/   Uint4B                       Flags
/0x020/   Uint4B                       Characteristics
/0x024/   Ptr32 _VPB                   Vpb
/0x028/   Ptr32 Void                   DeviceExtension
/0x02c/   Uint4B                       DeviceType
/0x030/   Char                         StackSize
/0x034/   _unnamed                     Queue
/0x05c/   Uint4B                       AlignmentRequirement
/0x060/   _KDEVICE_QUEUE               DeviceQueue
/0x074/   _KDPC                        Dpc
/0x094/   Uint4B                       ActiveThreadCount
/0x098/   Ptr32 Void                   SecurityDescriptor
/0x09c/   _KEVENT                      DeviceLock
/0x0ac/   Uint2B                       SpectorSize
/0x0ae/   Uint2B                       Spare1
/0x0b0/   Ptr32 _DEVOBJ_EXTENSION      DeviceObjectExtension
/0x0b4/   Ptr32 void                   Reserved
}
Object_Device     *pObject_Device
```

Fig. 9. Definition of an Object Device

F. Removing Drivers from PsLoadedModuleList

A PsLoadedModuleList structure whose definition is shown in Fig. 10 can be found from Fuzun, FU rootkit [8] writer, who hid drivers through it. In this trick, we can check our loaded rootkit driver to get its Object Driver's DriverSection, and further find PsLoadedModuleList so that we can traverse it to get the desired driver addresses to hide.

```
typedef  struct  PsLoadedModuleList
{
/0x000/   _List_Entry                  TypeList
/0x008/   Ptr64                        UnKnow
/0x018/   Ptr32 Void                   DriverStart
/0x01c/   Ptr32 Void                   DriverInit
/0x020/   Uint4B                       UnKnow1
/0x024/   Ptr32 _UNICODE_STRING        Driver_Path
/0x028/   Ptr32 _UNICODE_STRING        Driver_Name
}
 PsLoadedModuleList    * pPsLoadedModuleList
```

Fig. 10. Definition of a PsLoadedModuleList

G. Altering Object Driver's Appearance

The targeted Object Driver appearance is modified to let it look different as compared to a normal one. This method tries to escape signature-based detectors. For example, if the value stored in the offset 0x000h of an Object Driver should be 0x04, then we can alter it with a random value to accomplish the purpose of stealth.

4 Experimental result and analysis

In Section 3 we have depicted the seven stealth technologies of the proposed new Windows-based rootkit in cloud operating systems. In the following, we will demonstrate the experimental results of testing the proposed rootkit's stealth ability in Windows cloud operating systems, which can be divided into the following two phases: (1) rootkit loading operations, (2) test and analysis of rootkit stealth ability.

4.1 Rootkit loading operations

The proposed rootkit named rootkit.sys is a driver format and executed on Windows server 2008 R2 cloud computing service where multiple virtual machines are co-located on the same physical server. In such systems, physical resources are transparently shared by the virtual machines belonging to multiple users. For the rootkit loading operations, the proposed rootkit is installed but its stealth functionality is not invoked, as shown in Figs. 11 and 12. In Fig. 12, we can clearly find the rootkit driver name "rootkit.sys" after it is loaded, and therefore it attests that the proposed rootkit is successfully loaded into the cloud operating systems.

4.2 Test and analysis of rootkit stealth ability

In order to provide greater flexibility in the testing process, the GUI (Graphical User Interface) interface is designed to test the stealth ability of the proposed rootkit, as shown in Fig. 13. In Fig. 13, since we have implemented seven hiding techniques in the proposed "rootkit.sys" drivers, testers can

check off their required needs of the concealment mechanism to deploy a variety of different type of rootkit.

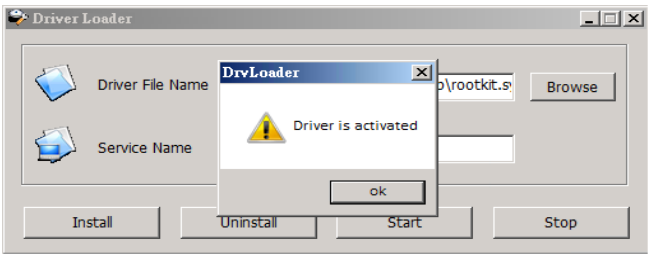


Fig. 11. The proposed rootkit "rootkit.sys" is installed

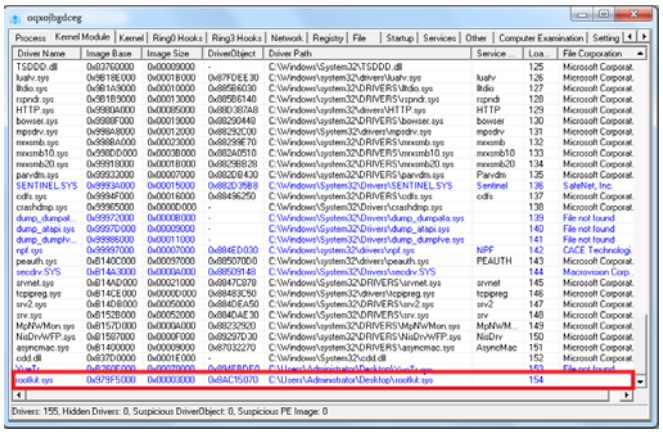


Fig. 12. The stealth functionality of the proposed rootkit "rootkit.sys" is not invoked, and therefore it is shown in the cloud operating systems

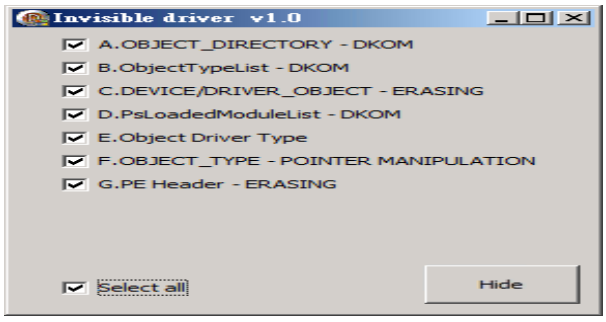


Fig. 13. The GUI for testing the stealth functionality

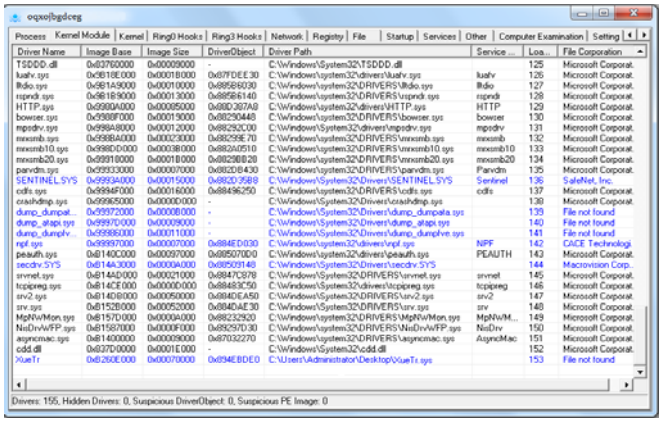


Fig. 14. After the stealth functionality of the proposed rootkit "rootkit.sys" is invoked, it disappears in the cloud operating systems

When the seven stealth tricks of the proposed rootkit "rootkit.sys" are checked off in Fig. 13, "rootkit.sys" disappears in the cloud operating systems, as shown in Fig. 14, and thus proves that it have launched the hiding feature. According to the recommendations of the ICSA Labs's website [24], we find the three anti-rootkit tools developed by AVAST, Trend Micro, and ESET respectively can be effectively used for rootkit detection in cloud computing environments. Therefore, in this paper the aforementioned three rootkit detectors are used to test the stealth ability of the proposed rootkit. When the stealth functionality of the proposed rootkit has been invoked, all of the three detectors cannot detect the presence of the proposed rootkit in cloud computing environments, as shown in Figs. 15-17.

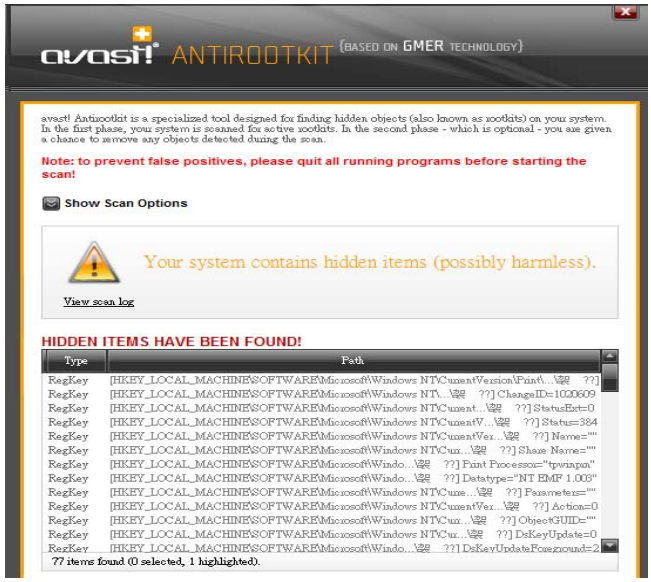


Fig. 15. AVAST anti-rootkit tool cannot detect the proposed rootkit

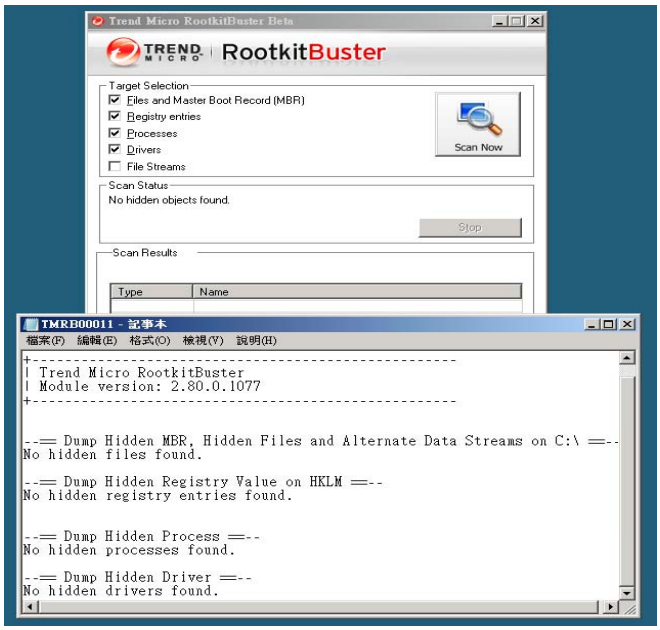


Fig. 16. Trend Micro anti-rootkit tool cannot detect the proposed rootkit

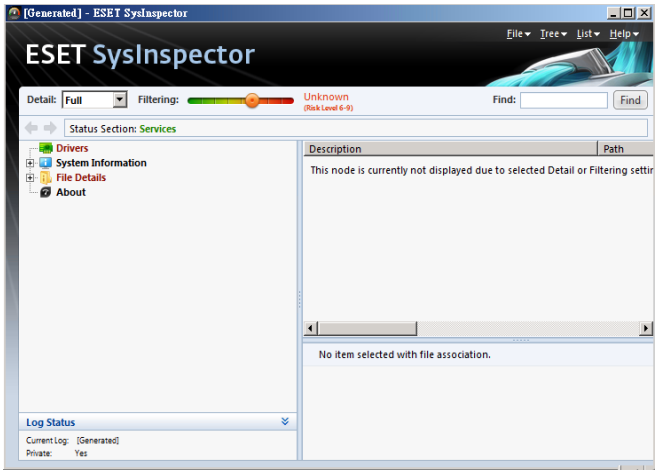


Fig. 17. ESET anti-rootkit tool cannot detect the proposed rootkit

We conclude that why all of the tested detectors cannot detect the proposed rootkit driver should be the following reasons. One is that some detectors cannot detect the rootkit with the abilities of removing (1) the signature of PE image and (2) the driver-related information in the registry. Another is that some detectors employ memory scan with predefined signatures, but they cannot recognize hidden Object Drivers with an abnormal object appearance. The other is that some detectors do not completely check whether the List_Entry data structures of Object Directory, Object Driver, Object Device and PsLoadedModuleList may be modified, and thus the rootkit with the tricks of modifying the List_Entry data structures can avoid the heuristic-based detectors.

5 Conclusion

In this paper, we come up with new Windows rootkit based technologies for enhancing the DRM in preventing the confidential information of digital contents from unauthorized usages by illegal users in cloud computing environments. The proposed new driver-hidden rootkit executed on Windows cloud operating systems has successfully evaded the well-known anti-rootkit detectors, and thus can be effectively used to prevent unauthorized users from removing the rootkit of concealing the DRM software by employing anti-rootkit tools in cloud computing environments.

To the best of our knowledge, there are no literature exploring the rootkit-based technologies for enhancing the DRM in cloud computing environments at present, so this paper is the first attempt to develop rootkit-based protection technologies against unauthorized usages of digital contents in cloud computing environments. We affirm that in cloud computing environments our research is valuable for extending the protection of the DRM software, and can be a great inspiration to DRM software makers to effectively improve the current techniques of defending against the illegal distribution and consumption of copyrighted digital multimedia contents. Furthermore, this study also inspires defenders to effectively strengthen the legitimate uses in cloud computing environments by the hiding tricks of the proposed subtle rootkit.

Acknowledgement

This work was supported by the National Science Council of Republic of China under contract numbers NSC 102-2219-E-212-001 and NSC 101-2221-E-212-006-MY3.

References

[1] E. W. Felten and J. A. Halderman, "Digital rights management, spyware, and security," *IEEE Security & Privacy*, vol. 4, no. 1, pp.18-23, 2006.

[2] S. Lian, "Digital rights management for the home TV based on scalable video coding," *IEEE Trans. Consumer Electron.*, vol. 54, no. 3, pp. 1287-1293, Aug. 2008.

[3] Y. Zou, T. Huang, W. Gao, and L. Huo, "H.264 video encryption scheme adaptive to DRM," *IEEE Trans. Consumer Electron.*, vol. 52, no. 4, pp. 1289-1297, Nov. 2006.

[4] S. Park, J. Jeong and T. Kwon, "Contents distribution system based on MPEG-4 ISMACryp in IP set-top box environments," *IEEE Trans. Consumer Electron.*, vol. 52, no. 2, pp. 660-668, May 2006.

[5] G. Kim, D. Shin, and D. Shin, "An efficient methodology for multimedia digital rights management on mobile handset," *IEEE Trans. Consumer Electron.*, vol. 50, no. 4, pp. 1130-1134, Nov. 2004.

[6] P. Zou, C. Wang, Z. Liu, and D. Bao, "Phosphor: A cloud based DRM scheme with sim card," *Proceedings of 12th International Asia-Pacific Web Conference*, pp. 459-463, 2010.

[7] W. Tsaur, "Strengthening digital rights management using a new driver-hidden rootkit," *IEEE Trans. Consumer Electron.*, vol. 58, no. 2, pp. 479-483, May 2012.

[8] L. Stevenson and N. Altholz, Rootkits for Dummies, Wiley Publishing, 2007.

[9] E. U. Kumar, "Battle with the unseen — understanding rootkits," *Proceedings of the 9th AVAR International Conference*, pp. 82-97, 2006.

[10] C. Keong, "Defeating kernel native API hookers by direct service dispatch table restoration," *Technical Report*, SIG2 G-TEC Lab, October 2004.

[11] G. Hunt and D. Brubacher, "Detours: binary interception of functions," *Proceedings of the Third USENIX Symposium*, pp. 135-143, 1999.

[12] P. Beaucamps, "Advanced polymorphic techniques," *International Journal of Computer Science*, vol. 2, no. 3, pp. 194-205, 2007.

[13] J. Bulter, J. L. Undercoffer and J. Pinkston, "Hidden process: the implication for intrusion detection," *Proceedings of the IEEE International Workshop on Information Assurance*, pp. 116-121, 2003.

[14] A. Baliga, L. Iftode, X. Chen, "Automated containment of rootkits attacks," *Computers & Security*, vol. 27, no. 7-8, pp. 323-334, 2008.

[15] C. Xuan, J. Copeland and R. Beyah, "Shepherding loadable kernel modules through on-demand emulation," *Proceedings of the 6th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pp. 48–67, 2009.

[16] D. Lobo, P. Watters and X. Wu, "Identifying rootkit infections using data mining," *Proceedings of 2010 International Conference on Information Science and Applications*, pp. 1-7, 2010.

[17] D. Lobo, P. Watters and X. Wu, "RBACS: rootkit behavioral analysis and classification system," *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining*, pp. 75-80, 2010.

[18] A. Baliga, V. Ganapathy and L. Iftode, "Detecting Kernel-Level Rootkits Using Data Structure Invariants," *IEEE Trans. Dependable and Secure Computing*, vol. 8, no. 5, pp. 670-684, Sept. 2011.

[19] J. Li, Z. Wang, T. Bletsch, D. Srinivasan, M. Grace and X. Jiang, "Comprehensive and Efficient Protection of Kernel Control Data," *IEEE Trans. Information Forensics and Security*, vol. 6, no. 4, pp. 1404 - 1417, Dec. 2011.

[20] A. Baliga, P. Kamat and L. Iftode, "Lurking in the shadows: identifying systemic threats to kernel data," *Proceedings of the 2007 IEEE Symposium on Security and Privacy*, pp. 246-251, 2007.

[21] C. Kreibich and J. Crowcroft, "Honeycomb: creating intrusion detection signatures using honypots," *ACM SIGCOMM Computer Communication Review*, pp. 51-56, 2004.

[22] N. L. Petroni Jr., T. Fraser, A. Walters and W. Arbaugh, "An architecture for specification-based detection of semantic integrity violations in kernel dynamic data," *Proceedings of the 15th USENIX Security Symposium*, pp. 289-304, 2006.

[23] Y. Ye, D. Wang, T. Li, D. Ye and Q. Jiang, " An intelligent PE-malware detection system based on association mining," *Journal in Computer Virology*, vol. 4, no. 4, pp. 323-334, 2008.

[24] ICSA Labs, http://www.icsalabs.com, 2014.

On the Automated Passengers' Authenticated Bus Transit Services in South Africa

Bassey Isong, Egbe Dominic, Gilbert Dzawo

Department of Computer Science and Information Systems, University of Venda
Thohoyandou, South Africa

Abstract *— Today, the impact of information technology has been beneficial and witnessed in every key sector of human endeavour such as healthcare, transportation, education, e-government, entertainment, nations and so on. However, in the perspective of transportation, the existing bus transit system in South Africa (SA) is still characterized by inefficient manual operations such as poor quality of service, non-user friendly and inefficient passenger identification. In this case, due to the economic and developmental role played by the bus transport sector in SA, it is important that some advanced information and communication technology (ICT) tools be incorporated into its operations. This requires a system that automates all the bus transit services and has the capability to authenticate passengers in a real-time mode. Therefore, this paper, proposes an automated bus transit system that offers e-subscription and real-time biometric passengers' authentication. This will proffer solution to the challenges faced by the bus transit industry in SA. Furthermore, a novel system prototype is developed with the technologies required to fill the service gap.*

Keywords *– Transportation, Real-time, Biometric, Bus*

I. INTRODUCTION

Transportation plays a critical role worldwide and has remain one of the most vital and indispensable organ of human activities. Ranging from the most ancient to our modern forms of transportation such as aircrafts, ships, speed trains, and buses or cars, the value of transportation still maintains its high significance to human and economic activities. In this case, the societies cannot function without transportation. According to [1], transportation importance is described as "boarding a train or plane, or dispatching freight to the other side of the world are, for hundreds of millions of people, for hundreds of thousands of businesses, everyday acts which are both unremarkable and vital".

Transportation is defined as "the movement of people, animals and goods or services from one location to another" [2]. The overall importance of transportation cannot be overemphasized because it facilitates the establishment of physical contact between people, goods, and services thereby promoting economic and social life. Like many economic activities that are intensive in infrastructures, the transport sector is an important component of the economy impacting on development and the welfare of populations. An efficient transport system has the potential to provide economic and social opportunities as well as benefits that result in positive multipliers effects such as better accessibility to markets, employment and additional investments [3]. By implication, an efficient system of transport is essential for a modern society with a competitive economy like South Africa (SA) for the purpose of rapid development. The current structure of SA transport system is divided into ports and shipping, roads, railways, airports and airlines. According to [5], "available statistics indicate that 80 per cent of SA's population is totally dependent on public transport (bus, commuter rail and taxis) for its mobility needs". But unlike other means of transport, the bus transport system is yet to fully incorporate information and communication technology (ICT) tools into its services. These current bus transit companies still run all their operations ranging from ticket subscription to passenger authentication using the traditional manual approach.

Meanwhile automation plays an increasing important role in the world's economy and in daily experience. Transportation is not an exception as it is quite evident in modern day information technology (IT) enabled air ticket booking and reservation systems as well as rail and bus ticketing systems such as eFare 2010, MyWay smart card system, Gautrain etc. All these are already being implemented in several countries to help improve the quality of service in the transportation industry. Also running our mass transit systems manually would equally imply the absence of a central point of coordination. This is a situation where individual transport companies adopt isolated business and administrative policies. It is definitely not too healthy for both clients and the government since it introduces two important challenges; firstly, it makes it difficult for the government to monitor transport fares or implement a uniform transport fare system for equivalent distances across the country. Secondly, there is the huge challenge of getting accurate data for appropriate budgeting, planning and forecasting in the transport sector. In addition, other issues related to improving the quality of customer service and ensuring efficient passenger authentication to control fraud also remain a major problem. Therefore, in this research we propose an automated passenger authentication bus transit system that will be able to store passengers' data in a central database and handle authentication of passengers in real-time mode using fingerprint technology. The goal is to ensure that the bus transit system operates effectively and efficiently in SA cities and is beneficial to users and operators. As a proof of concept, we developed a system prototype called Automated Passengers Authenticated Bus Transit System (APABTS) to simulate the system's operations.

The rest of the paper is organized as follows: Section II is the current state of bus transit in SA, Section III is the design

and analysis of the system, Section IV is the prototype implementation, Section V APABTS operation, Section VI discussion and Section VII is the paper conclusion.

II. Current State of the Bus Transit System

The SA land transport system presents a scenario that will be quite suitable for automation because it is dominated by several privately run intra-city bus transit systems, offering monthly services on fixed routes to customers such as school children, students, civil servants, and others. To use these bus services, customers (passengers) have to pay a monthly fare or service subscription that covers their daily to-and-fro movement on a fixed bus route throughout the month subscribed for. But as mention earlier, the bus transit system is still being mired with series of challenges owing to its manual based operations. Interestingly, it has been revealed that a high proportion of average SA's students, businessmen, private, and government workers depend on the private bus mass transits for mobility [5]. Moreover, a significant amount of economic activities hinges on transportation service, making it obvious that an effective and efficient transportation system will no doubt have a direct positive economic impact or otherwise. Transport is often described as the heartbeat of SA's economic growth and social development [8]. Also, direct and indirect job creation in SA has been linked to transport and its related services, being a catalyst for economic growth [9].

The above fact in itself establishes the weight of the necessity to consider the option of an automated intra-city mass transit solution that will utilise the power of ICT to provide automated services. This paper will potentially address the concerns about the manual bus system through an automated passenger authentication bus transit system that will be web-based, support parallel ticketing, and used fingerprint to authenticate passenger in real-time mode.

III. Proposed System Analysis and Design

A. Functional Requirements

For efficient and professional system development, there is need to first adequately elucidate the system requirements that will serve as blueprint for system design and development. This section then outline the major system requirements namely – basic, general, passenger subscription, and passenger identification requirements.

1) Basic Functional Requirements: The requirements that form the core of the entire system are as follows:

R1-1: The system shall only allow a passenger to register only once on the subscription portal. - *Integrity*
R1-2: The system shall allow only VALID passengers to use the bus – *Authentication*

2) General Requirements: The general requirements take care of system usage and privileges. That is, those who can use the system and at what level of allowable privilege. These include three kinds of system users – passenger, system administrator and support staff. It is here required that only authenticated persons are able to access or make use of the system and in accordance with their level assigned privileges. It includes the following two requirements:

R2-1 – Sign Up: This is the aspect of the system that enables certified system users to choose the platform of privileges to be associated with throughout the period of their system use by creating an account on the portal. This requirement ensures that only persons with certified credential are allowed to gain access to or use the system and that such a system user's tasks are only restricted within the boundaries defined by the associated privileges.

R2-2 – Sign In: With this feature authentic users are allowed to gain access to or use the system. It is meant to guarantee that only certified passengers, system administrator or support staff are enabled to use the system to perform tasks as defined by the associated privileges.

3) Passenger Subscriptions: Since the system makes use of an online portal and database that stores both passenger and administrative information for security, accounting, and support purposes, there has to be some features to support such functionalities as mentioned above. These are expressed in the following requirements:

R31 - Ticket Subscription: The ticket subscription feature enables a passenger to pay and subscript for a bus ticket through the online portal. With this it becomes possible for passengers to for bus ticket individually and online, run separate subscription periods independent of others.

R3-2 - Data capture: This feature enables the system administrator or support staff to capture passengers' fingerprint. This is meant to create fingerprint authentication functionality in order to guarantee the one-passenger-one-subscription objective.

R3-3 – Renew Subscription: This feature allows users to renew their subscription after its expiration period. With this there will be no need for passengers to physically visit the bus ticket office for new subscription, making the system more user-friendly.

4) Passenger identification: The major challenge with the manual bus ticketing system is the difficulty in ascertaining the authenticity of passengers and the ease with which hard copy cards can be swapped. Passenger identification is handled differently with the requirement defined below:

R4-1 - Validation: This feature enables a passenger to gain access into the bus by authenticating his/her identity via a fingerprint device. It ensures only passengers with a matching finger print in the database are able to use the bus facility.

B. Use Case Analysis

As part of the requirements analysis and modeling, this section presents the representation of the system actors and their roles for effective understanding of the overall system processes.

1) Actors: Here three types of system users are identified as actors: *the passengers, the support staff* and *system administrator*. These are users who are going to interact with the system in order to achieve its objectives. This is captured in Figure 1.

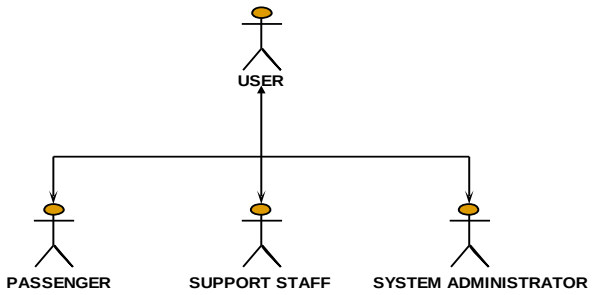


Figure 1. System Actors

2) Use Case Model: The roles of the three major system actors of APABTS (passengers, support staff, and system administrator) are clearly demonstrated here, showing interactions between system actors and system functionalities in a general use case model.

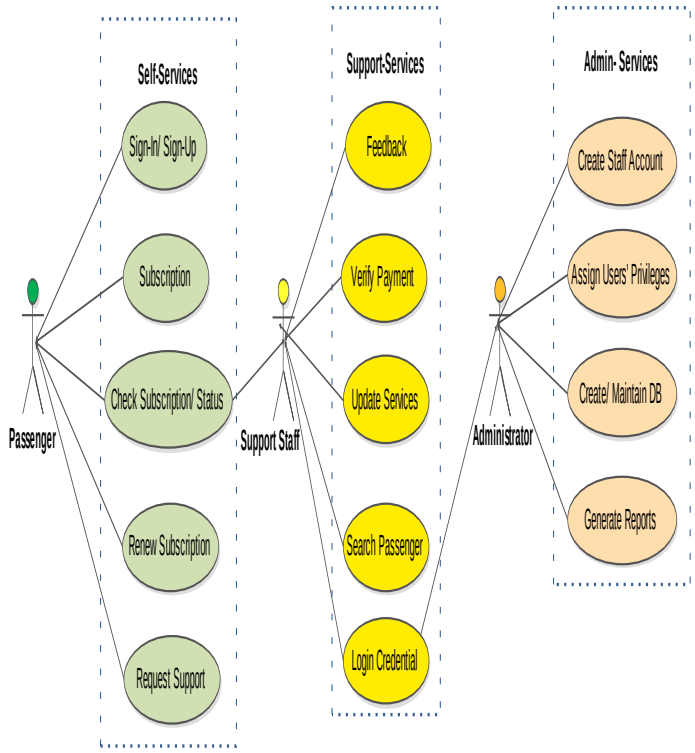


Figure 2. System use case model showing actors and interactions

C. System Structure and Components

The automated passenger authentication bus transit system requires a number of components to achieve its e-fare subscription and passenger authentication objectives as discussed section 3.1. These components include the following:

1) Biometric Sub-systems: The biometric sub-system is one of the most important components of this proposed system because it plays the central role in passenger authentication. It consists of a fingerprint scanner/reader or sensor with an inbuilt database (made to interface with another database on a remote server) and can be set to either enrolment or authentication mode. A good example of such biometric system is the FingerTec technology, with 360° rotation of live capture for identification or verification, high identification speed, small memory requirement (where whole matching algorithm required just 350KB). It is an online biometric fuzzy inference system (OFIS) solution used for real-time verification and enrolment. The FingerTec OFIS runs on Browser/Server (B/S) Environment, where users can enrol their fingerprint through the FingerTec OFIS Scanner that is linked to a PC [10]. We strongly recommend FingerTech biometric technology for the implementation of this proposed system due to its compatibility and performance efficiency.

2) APABTS Database: One of the key features of the proposed system is the use of a secure online data repository to hold passengers' information. Making use of a database is highly significant in the sense that the automated FIS references it in the process of searching for a match against individual passenger's fingerprint. Also, it helps to make transportation data handy for the purpose of forecasting and planning. The proposed structure is captured in Figure 3.

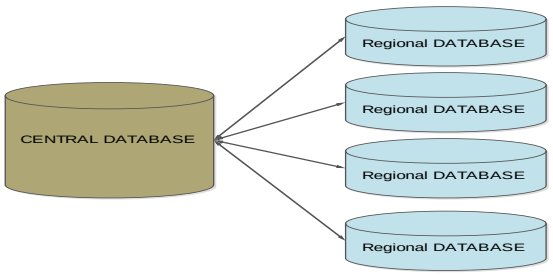


Figure 3. Proposed database structure

Since the proposed system will be used by different bus transit companies, the database system will be structured to consist of a central database (holding data from all bus transit companies) and regional databases that will store data from each bus transit company. With this database architecture, system administrators will only have access to their bus transit database alone, whereas all the regional databases are linked up to the central database that will be only accessed by government or authorized government agencies for the purpose of monitoring, forecasting, and planning. The database will hold passenger information such as *name, fingerprint, gender, passenger ID, e-mail, phone number*, and *city of residence*. In addition to these passenger personal data, the subscription portal generates transaction data for each passenger as he or she interacts with the e-portal. The transaction information that will be generated for each passenger includes the following:

1) Subscription status
2) Subscription date
3) Expiration date, and
4) Payment status.

In the database, information will be organized in three tables, namely ***personal data table*** (for passenger information including fingerprint data), ***payment table*** (to contain passenger payment details), and ***subscription table*** (for data fields which are generated as passengers interact with the portal). Information from these tables is to be used by the system administrator for support purposes and informed decision making.

3) Supportive Devices: Apart from the core system components as explained above, other devices are also needed to complement the running of the system such as Internet facility. The processes of e-subscription and automated passenger authentication functionalities of the proposed system are both web-based and therefore, require a fast and reliable internet access point wired or wireless. For the purpose of mobility, we recommend any mobile ultra-broadband Internet access device such as wireless 4G LTE mobile hotspot MiFi 4510L in the buses for the fingerprint sensor for real-time passenger authentication.

In addition, laptops or desktop PCs constitute another important supportive device. For the bus transit companies to fully utilize the system for administrative purposes and also effectively provide the required services to their clients, it is highly recommended that the various bus service offices be equipped with these devices. Such computer systems should be up-to-date in terms their capability to support wireless internet facilities.

The conceptual model that defines the system structure and interaction between systems components discussed in this section is captured in Figure 4.

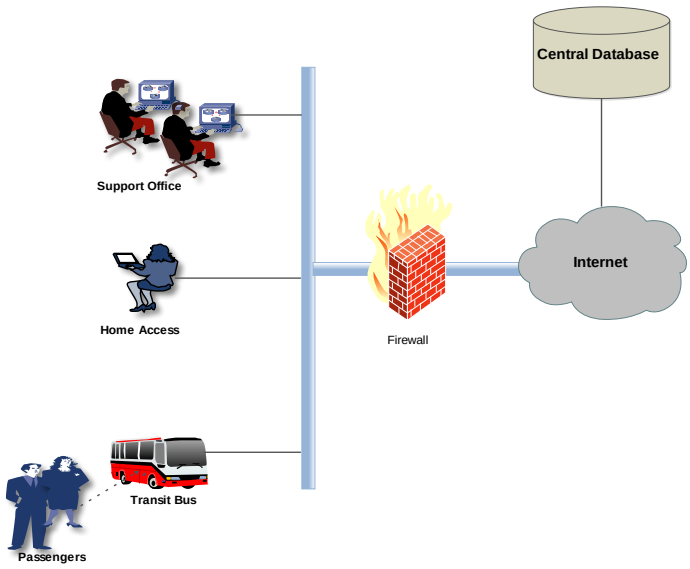


Figure 4. Proposed APABTS model

IV. APABTS Prototype Implementation

The core functionalities of APABTS partition the entire system into two interfaces: *admin* and *client interface*, each with a set of functionalities. The prototype implementation explain here is thus, based on these two partitions of the proposed system.

A. Admin Interface

The admin interface offers two types of login classification – supper admin and support staff. The supper admin account is configured once during the system deployment, after which the supper admin can create support staff account.

1) Supper Admin Functionalities: With supper admin credentials, an administrator performs the following tasks - add administrators, view subscription requests, assign privileges, treat client's enquiries, update bus fares.

B. Clients 'Interface

The admin interface offers two types of login classification – *supper admin* and *support staff*. The supper admin account is configured once at during system deployment then the supper admin can subsequently create support staff account. In compliance with the standard security principles of verification, validation, authentication, and authorization, APABTS provides this interface constrained for only clients with valid login credentials.

1) Client's Interface Functionalities: The client interface offers a variety of functionalities through which the bus transit services are accessed. These include: Sign Up, Login, view personal Details, order for subscription, read feed backs, check bus Fares.

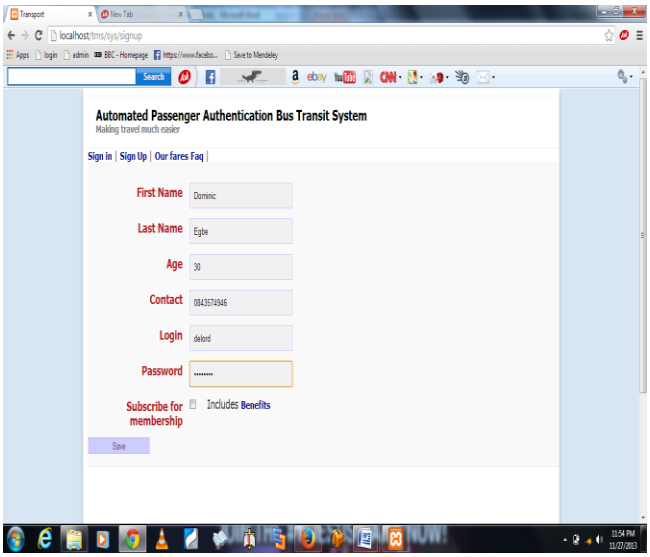


Figure 5. APABTS sign-up/sign-in page

C. APABTS Operations

In this section, we give a step by step explanation of the operation of APABTS system. These are presented as follows:

1) Sign up/sign in: This functionality enables a user to create account in the subscription portal by supplying user name/password and other personal data required to maintain the client's profile on the database. (See Figure 5)With the chosen username and password the user can login if account creation is successful.

2) Subscription Request: In order to request for subscription, the user can click on "**My subscription**" (after successful login), selects subscription plan, supplies payment details and sends a subscription request to the system administrator by clicking "**Subscribe**" button as captured in Figure 6.

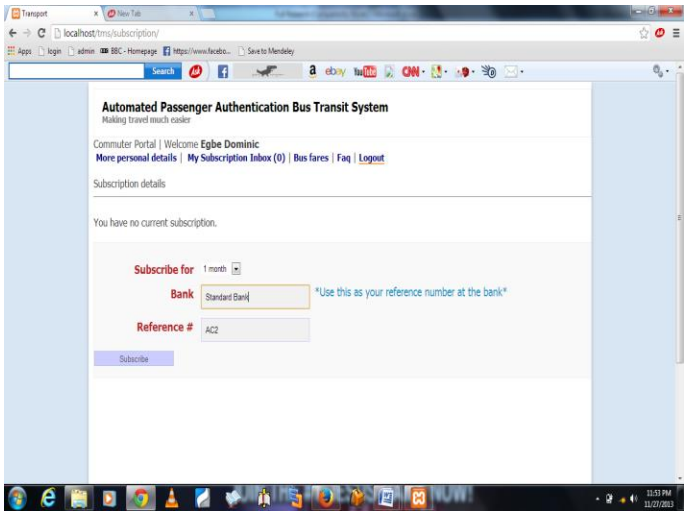


Figure 6. APABTS subscription request page

3) Subscription Approval: It is the duty of the supper admin or support staff (depending on the privilege level granted) to approve client's subscriptions from the admin system interface. (see Figure 7) To view all clients' subscription requests, the system administrator clicks on "**Subscription requests**" and selects on "**Full details**" option to activate individual client request as shown the figure that follows.

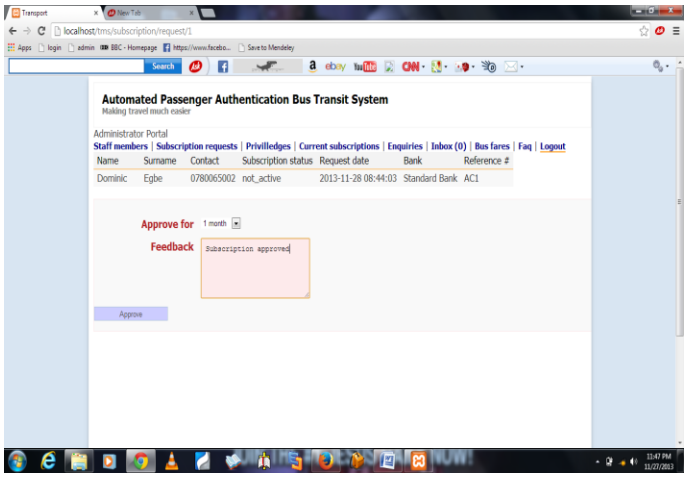


Figure 7. APABTS subscription status page

4) View Subscription Status: Subsequently after submitting a request for subscription, the status of the request can be checked by users. To do so, a user can login and click on "My subscription". If it has been approved, the system displays the duration of the subscription otherwise, the message "Your subscription is awaiting approval" displayed. (see Figure 7)

V. PASSENGER IDENTIFICATION IN APABTS

One of the major requirements of the proposed system is to ensure one-client-one-subscription. To achieve this, the APABTS employs a biometric system using fingerprint to authenticate clients in a real-time mode. The whole process of passenger authentication is in twofold: enrolment and authentication.

A. Enrolments

This is when passenger fingerprint data is extracted and stored in the APABTS database using a biometric scanner. When first time clients create user account or login or send subscription request to the system administrator, they are prompted by the system to visit any of the support offices for biometric data capture. At the bus support office, the system administrator scans client's fingerprint and uploads it to the biometric database. On the biometric database, each client's fingerprint data is link to their unique ID number on the subscription table and personal data table.

B. Authentication

Here passengers are identified in real-time mode at bus stations by APABTS as they place their finger on the scanner devise installed beside the door of the bus as the walk in. During this process, the system verifies a match between their fingerprint and the image stored in the fingerprint database. The biometric scanner to be used for example FingerTech has a programmable interface that can be programmed in the implementation language to light the green indicator light if there is a valid match otherwise the red indicator light.

VI. DISCUSSIONS

For any competitive economy, the need to have an effective and efficient transportation system cannot be over emphasized. The emerging SA economy that also plays a leading role in Africa is not an exception. Having a well efficient air and rail systems, SA bus transit system is yet to braze up with the information technology demands of the 21st century. The fact remains that the transport sector just like several other economic activities that are intensive in infrastructures, is an economic catalyst, impacting on development and the welfare of populations. Therefore, having an efficient transport system in place has the potential to provide economic and social opportunities as well as benefits such as better accessibility to markets, employment and additional investments [3]. It must be noted that any transport system is service oriented in nature and the core parameter with which to define its efficiency is the quality of service rendered.

But opposed to this background, we still run a bus transit system in SA characterized by poor quality of service and

fraud. These elements of inefficiency in the bus transit system, to a large extent stem from the lack of ICT-driven operations in the system unlike what is obtainable in other systems of transportation – air, rail. It is against this backdrop that we are proposing an APABTS, an IT-driven bus transit system for fixed route buses. The design of APABTS is aimed at eliminating the traditional manual card transaction system, replacing it with an e-subscription portal with biometric capability to perform passenger authentication. This approach will bring several benefits in addition to minimizing fraud and improving the quality of services. First, it will bring about increased revenue generation for bus transit companies. Second, improving the quality of service and boosting revenue generation in the bus transit system will attract more investors thereby raising job creation opportunities. Finally, the architecture of APABTS makes provision for a central, secured remote database that can serve as a dependable data repository for SA department of transport for the purpose of budgeting and planning.

VII. CONCLUSION

Transportation is a key component of any competitive economy apart from being an agent of job creation and a catalyst to social development. And the world is fast incorporating essential ICT to every facet of human endeavour to improve productivity and quality of service amongst a host of other benefits – the transport sector is one of such major benefactors. The state of affairs of bus transport system in SA places a demand for the full incorporation of IT in order to boost its operations and deliver quality of service to its huge consumers. In this paper, we have offered the background justifying the need for automation in SA bus transit sector, and present the system requirements, design and implementation of a novel bus transit system, APABTS for commercial bus companies with fixed routes. The system consists of an e-portal with which passengers are able to subscribe for bus service online and a biometric component with which passengers are authenticated at bus stations real-time using fingerprint. Based on the components incorporated into this system, it is believed that if implemented for use, APABTS will be more user-friendly, flexible, secured and fault tolerant than its counter parts – manual and other existing systems.

REFERENCES

[1] E. Commission, "World report on road traffic injury prevention," *World Rep. Road Trafic*, 2012.

[2] N. Encyclopedia, "Transport - New World Encyclopedia," 2010. [Online]. Available: http://www.newworldencyclopedia.org/entry/transport. [Accessed: 10-Aug-2013].

[3] Rodrigue and T. Notteboom, "Transportation and economic development," 2012.

[4] "The state of South Africa's public transport system | earthworks magazine." [Online]. Available: http://earthworksmagazine.co.za/features/1452-2/. [Accessed: 19-Aug-2013].

[5] "SABOA, Industry Profile, Southern African Bus Operators Association, Public Bus Transport." [Online]. Available: http://www.saboa.co.za/industry_profile.htm.

[6] A. Systems and O. Platform, "Automation Systems the Optimisation Platform," 2013.

[7] Ouellette, R. *Automation Impacts on Industry*. Ann Arbor, MI, USA: Ann Arbor Science Publishers, 1983.

[8] "Department of Transport - South Africa > Home." [Online]. Available: http://www.transport.gov.za/. [Accessed: 15-Aug-2013].

[9] "South Africa Government Online." [Online]. Available: http://www.gov.za/.

[10] B. Solutions, C. Fingerprint, P. Free, E. Integration, M. Y. Own, A. Securely, P. Recognition, and I. Programming, "FingerTec ' s Online Fingerprint Identification System or OFIS is a biometric solution for online verification and enrollment . The FingerTec."

A Study of Poster and Viewer Participation in SNS

Shwu-Min Horng
Department of Business Administration, National Chengchi University, Taipei City, Taiwan (ROC)

Abstract - *This article studies the individual and social factors influencing the participation intention of viewers and posters in virtual communities. Reward, structural capital, and trust play a significant role for both types of user. Reciprocity has a positive impact only on the viewers' participation intention, while posters are also affected by reputation and cognitive capital. The positive relationships between cognitive capital and participant intention, and between trust and participation intention are stronger for posters than for viewers. In addition, managerial implications and suggestions for future research are discussed.*

Keywords: Virtual communities, SNS, social capital, user participation, viewers and posters

1 Introduction

The internet-based social network sites (hereinafter referred to as SNS) has drawn numerous attention from researchers as well as practitioners and its importance can be signified by their ever increasing usage [1]. SNSs enable users to communicate and connect with each other, to build up a personal network with common interests, allowing them to interact regularly in an organized way over the Internet. Although it is not clearly defined, SNS can be considered equivalent to a virtual community in a broader sense [2].

The key development with respect to the virtual community is the great increase of user-generated content on the Web, and the ability to easily search through it and combine parts of it to form new content. Therefore, encouraging users to provide content becomes an important issue for a given virtual community to attract more users and sustain its competiveness [3]. If providing content may be termed a posting activity, another activity of viewing, along with the posting activity, is made up of the fundamental elements in the ongoing life of any virtual community [4]. Viewing or lurking has not received much attention and few studies on these activities are to be found in a review of the literature, since most research tends to focus on active participants, that is, those who post online.

Researcher had controversial opinions toward viewing. Viewers were labeled as “free-riders” who drain the social capital of the community because reading essentially means taking without giving back [5]. In contrast, another study presented lurking in a more positive light [6]. They discovered that many viewers considered themselves as community members, and were possessed of the characteristics that community members attribute to a successful online community [7].

Although viewers usually enjoy content on websites provided by others and do not actively participate in online communities, they account for the majority of users in many communities [8]. For online community sponsors and operators, viewers are important because they are part of the traffic, contributing to volume on servers, and respond to advertising and selling. It is important to understand how posters and viewers behave differently, so the online service operators can have strategies targeting their users more effectively.

2 Theoretical background

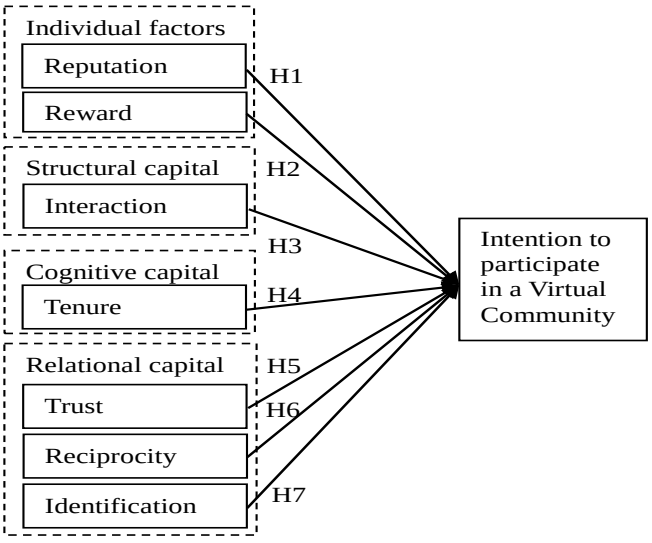


Fig. 1 The research model for virtual community participation

Fig. 1 shows the research model for the hypotheses in this study. The factors governing the intention of viewers and posters to participate in virtual communities are tested and compared. In the following sections I will discuss

each of the constructs and their relationships to virtual community participation.

2.1 Individual motivations

As discussed previously, participation in virtual communities includes passively viewing and actively posting, which was modeled as contribution of knowledge [9]. In order to share knowledge with others, individuals must deem their contribution to be worth the effort, thereby generating new value. They also expect to receive some of that value for themselves [10]. The cost and benefit factors based on social exchange theory explain human behavior in social exchanges [11]. An interaction is considered as being rewarding if the benefit perceived by the subject is greater than the effort experienced. This implies that an individual can benefit from active participation if he/she perceives that participation enhances his/her personal reputation in the network.

Although viewers do not interact with others and are considered free-riders [5], they are driven by the same factors affecting the posters who participate in SNS, if they are considered as the potential posters [12]. This inference leads to the first two sets of hypotheses.

H1a: Reputation is positively associated with a higher level of participation by viewers.
H1b: Reputation is positively associated with a higher level of participation by posters.

Another important individual motivation is the reward provided by a web service to encourage user participation. Rewards are considered one of the extrinsic forms of motivation intended to increase participation in virtual communities [13]. In practices, various forms of rewards are implemented for both of viewers and posters to encourage users' participation. Participation will be enhanced when the rewards of participating exceed the costs [14], which leads to the following hypotheses.

H2a: Rewards are positively associated with a higher level of participation by viewers.
H2b: Rewards are positively associated with a higher level of participation by posters.

2.2 Social capital theory

Researchers have examined the role of social capital in the creation of intellectual capital, and proposed that social capital, i.e., the network of relationships possessed by an individual and the set of resources embedded within it, strongly influences the extent to which interpersonal knowledge sharing occurs [10]. Social capital was defined in terms of three diverse dimensions, 1) structural capital represents links or connections between individuals; 2) the cognitive dimension focuses on the shared meaning and understanding that individuals or groups have with one another; and 3) relational capital refers to the personal relationships people have developed with each other through a history of interactions.

The connections between individuals, or the structural links created through the social interactions between individuals in a network, are important predictors of collective action [15]. This structural capital is the infrastructure of human capital that provides the environment which encourages individuals to create and leverage their knowledge by investing their human capital. Social interactions and ties were considered as channels for information and resource flows [16]. Individuals who occupy a central position within a collective have a relatively high proportion of direct ties to other members, and are likely to have developed the habit of cooperation. Moreover, such individuals are more likely to comply with group norms, which in turn lead to participation in a virtual community [17]. Although viewers are not the center of the virtual community, they could still build their connections with posters and other viewers through the center of the virtual groups. Thus, the following sets of hypotheses are:

H3a: Structural capital is positively associated with a higher level of participation by viewers.
H3b: Structural capital is positively associated with a higher level of participation by posters.

Cognitive capital represents the shared meaning and understanding that individuals or groups have with one another. Shared language and vocabulary influence the conditions for knowledge exchange in several ways [10]. As an individual interacts over time with others by sharing knowledge and learning the norms of practices in a virtual community, that individual develops his/her cognitive capital. Cognitive capital consists of mastering the application of expertise, which takes experience to build [9]. Individuals with a longer tenure in shared practice are likely to better understand how their expertise is relevant and may thus be more highly motivated to share knowledge as well as utilize a knowledge-sharing mechanism [18]. From the discussions above, it appears that the shared meaning and understanding within a virtual community could exist among both of viewers and posters. This leads to the following hypotheses:

H4a: Higher cognitive capital is positively associated with a higher level of participation by viewers.
H4b: Higher cognitive capital is positively associated with a higher level of participation by posters.

Trust is a key element in establishing long-term customer relations in virtual communities [19], and it is an important factor in conducting transactions online [18]. The value of virtual communities to sponsoring firms is dependent on the sponsor's ability to cultivate trust with the members of communities, and provide theoretical contributions [20]. Since legal details cannot always be implemented, trust is an essential ingredient of long-term business engagements [21]. Furthermore, the lack of face-to-face contact in virtual communities increases the perceived risk of a relationship among users [22], and the risk of exposure of personal information. Users are constantly questioning the issue of security online, and trust is clearly the critical issue why posters or viewers continuously use the web services. Therefore I propose:

H5a: Trust is positively associated with a higher level of participation by viewers.
H5b: Trust is positively associated with a higher level of participation by posters.

Reciprocity has been regarded as a benefit for individuals engaging in social exchange [23], or as the belief that current contributions will lead to future requests for knowledge being met [11]. Norm of reciprocity is a highly productive component of social capital, and the frequently cited reason for participation in virtual communities. In another research, reciprocity was defined as one of the perceived extrinsic sources of motivation that has a positive effect on individual's use of a knowledge-sharing mechanism [13]. Although not clearly defined, the focuses of these literatures are posters who are motivated by reciprocity to share their knowledge online. However, viewers are potential posters and might be encouraged by the same reason. Therefore the following hypotheses have been developed accordingly:

H6a: Norm of reciprocity is positively associated with a higher level of participation by viewers.
H6b: Norm of reciprocity is positively associated with a higher level of participation by posters.

Identification is the process by means of which individuals see themselves as one with another person or group of people [10], and is one of the internal motivations for participation [24]. In this study, users identify themselves as members of a community and align their goals with those of the community as a whole. Identification is also referred as an individual's sense of belonging and a positive feeling toward a virtual community, which is similar to emotional attachment to the community [25]. Viewers and posters are both the members of the online communities. Thus, I hypothesize that both viewers and posters are motivated by their identity in a virtual community:

H7a: Identification is positively associated with a higher level of participation by viewers.
H7b: Identification is positively associated with a higher level of participation by posters.

3 Data collection

The data were collected from a website established in 2007. It was ranked in the top 20 among SNS in Taiwan [26]. A survey from the users of a single web service provides this study with a clear classification for the type of users, viewers and posters. A banner with a hyperlink connecting to a web survey was posted on its homepage. In order to collect enough data from posters, a mechanism was designed so that the banner would also be triggered when users are posting. The first page of the questionnaire explained the purpose of this survey and ensured confidentiality.

After excluding 165 invalid questionnaires for various reasons, a total of 364 valid samples were available for analysis, yielding an effective rate of 68.8 percent. The numbers of viewing and posting users are 220 and 144, respectively. No significant differences were found across the demographic information between viewers and posters upon visual inspection, and the profile of the sample basically matches expectations, with a higher proportion of male, student, and younger users.

4 Results and analysis

Multi-item, five-point Liker scale items were used to measure the constructs in the model. Scales were developed based on the review of the most relevant literature. To ensure that face validity, an iterative evaluation process was implemented so that each item used represented its definition without ambiguity [27]. To further validate the measurement model, convergent validity and discriminant validity were tested with results satisfying the requirements suggested. In sum, the results of content validity, convergent validity, and discriminant validity enable this study to proceed to estimations of the regression models.

For latent constructs where multiple items are available, they are combined into one indicator according to the partial disaggregation model [28]. In contrast to models where every item is a separate indicator, this yields models with fewer parameters to estimate, and better ratios of cases to parameters, while reducing measurement errors to a certain extent.

The intention to participate in a virtual community was tested for viewers and posters, respectively, using regression analysis. The results are shown in Table 1.

Table 1 Regression results

		Viewers		Poster	
		Stand. Coeff.	Sig.	Stand. Coeff.	Sig.
H1	Reputation	1.433*	0.087	2.387**	0.011
H2	Reward	3.645***	0.000	3.477***	0.000
H3	Structural	2.191**	0.017	2.287**	0.021
H4	Cognitive	-0.373	0.785	2.219**	0.032
H5	Trust	1.910**	0.039	3.057**	0.002
H6	Reciprocity	1.769*	0.068	1.883*	0.094
H7	Identification	0.488	0.874	0.356	0.748
	R^2	0.416		0.427	
	Adjusted R^2	0.382		0.406	
	F	17.483***	0.000	12.548***	0.000

***p < 0.01; **p < 0.05; *p < 0.1

Both of the regression models are significant at 0.01, and the differences between R-square and adjusted R-square indicates no over-fitting problem for the number of independent variables. Regression models for viewer and poster intentions explain 41.6 and 42.7 percent of the variance, respectively. Both of the R2 are significant at the 0.01 level. For viewers, H1a, H2a, H3a, H5a, and H6a are supported. Reputation, rewards, structural capital, trust, and reciprocity have a significantly positive effect on viewers' intention to participate in the virtual community. H4a and H7a are not supported. Cognitive capital and identification capital have no significant relationship with the participation behavior of viewers.

Variables showing significant effects on the posters' intention are reputation, reward, structural capital, cognitive capital, trust, and reciprocity in which support hypotheses of H1b, H2b, H3b, H4b, H5b, and H6b. H7b is not supported indicating that identification has insignificant relationships with the posters' intention to participate in the virtual community. Note that both viewers and posters behave in a similar way with respect to the variables of reputation, reward, structural capital, trust, reciprocity, and identification. Reputation, reward, structural capital, reciprocity, and trust are all positively significant while identification is insignificant. Cognitive capital has diverse results for viewers and posters.

5 Conclusions and discussion

The tests of the hypotheses are summarized in Table 2. H7 is the only hypothesis not supported for both of viewers and posters. The case company was established relatively recently, and it is very likely still at an early stage of development. Its users have not built a strong sense of belonging toward this virtual community. In addition, the identification construct is insignificant for both of viewers and posters. On the other hand, rewards are the most significant factor for both posters and viewers in their participating behavior, and it might serve as a salient motivator for knowledge contributors when the identification factor is weak [11]. Various types of reward system are commonly observed for online services at their early development stage.

Table 2 Summarized results of the hypotheses tests

Hypotheses		Results	
		Viewer (a)	Poster (b)
H1	Reputation is positively associated with a higher level of participation	Supported	Supported
H2	Rewards are positively associated with a higher level of participation	Supported	Supported
H3	Structural capital is positively associated with a higher level of participation	Supported	Supported
H4	Higher cognitive capital is positively associated with a higher level of participation	Not supported	Supported
H5	Trust is positively associated with a higher level of participation	Supported	Supported
H6	Norm of reciprocity is positively associated with a higher level of participation	Supported	Supported
H7	Identification is positively associated with a higher level of participation	Not supported	Not supported

In addition to individual factors of reward and reputation, structural capital, trust, and norm of reciprocity are significant factors related to the intention to participate regardless of whether users are viewers or posters. Trust is the most important factor for a successful virtual community from the perspective of members as well as operators [29], and it is also positively related to the quality of knowledge sharing [25]. Trust is an antecedent factor for participating in virtual brand communities [22] and plays the same important role for any user to participate in a website on a continuous basis. For structural capital, the results are similar to those of another study [9], despite the slightly different measures applied.

It is interesting to note that cognitive capital has inconsistent results for viewers and posters. Reputation and cognitive capital are significant predictors of participation for posters, consistent with prior research in online settings for knowledge contribution [9], but are insignificant for viewers. Viewers are not motivated by reputation, possibly because they are not easily identified by other users and many virtual communities, including funP, provide the mechanism to rank posters based on either their posting frequencies and/or the responses to their postings from other users. In general, users with a higher ranking are considered to have better reputations in that virtual community.

Contrary to expectations, reciprocity and cognitive capital do not have a significant impact on the intention to participate for either posters or viewers. When posters have a higher level of cognitive capital--indicating a longer tenure in this virtual community and being accustomed to certain patterns of communication with other users in this site--they are more willing to participate. Viewers have considerably less structural capital implies fewer interactions with others; hence, the factor of cognitive capital for viewers becomes irrelevant to their intention of participation. As for reciprocity, it affects the intention of viewers, but not posters, to participate. These results are very similar to those of another study [13] in which reciprocity had a significant effect on the intention to use knowledge-sharing mechanisms and an insignificant effect on knowledge-sharing intention. This finding seems to be consistent with the argument that relational capital may not develop in electronic networks due to a lack of shared history, high interdependence, frequent interactions, and co-presence [30].

5.1 Managerial implications

Web operators should continue their reward system because it is the best way to motivate both viewers and posters. Nevertheless, the reward system requires a significant amount of resources and the company will not be able to match pace with the increase in users if it is to grow exponentially. Strengthening the sense of community builds loyalty among users and encourages them to revisit the site on a continuous basis [31], thereby reinforcing purchasing behavior [32]. The weak identification according to the survey indicates that the users of this site still lack a sense of community. Offline activities which have a positive impact on user participation, and making use of multimedia are suggested means of elevating a sense of belonging [4]. The initiation of offline activities is another unresolved issue concerning which activities can be activated by web operators or users. The latter seems to comply with the principle of Web 2.0 and will help the community grow on its own without additional resources, but the former is better controlled and monitored by web management.

Viewers are not motivated by reputation, something normally reserved for posters. However, it is feasible to provide several levels of privilege for frequent viewers. The significance of structural capital to user intention to participate indicates that web services should continue to enhance their platform facilitating the interaction of users. The cognitive factor affects only posters' participation, not viewers'. To ensure long-lasting users, a web service needs to encourage viewers to become posters in order to become accustomed to the common communication pattern in a particular web service. Simplifying the posting process and differentiating the privileges designated exclusively for posters should provide the incentives needed for this purpose.

Trust is the essential element of a web service and the value of a virtual community is dependent upon how the web operator is able to cultivate trust among the community's users [20]. Site management should be very careful when handling the personal information of their members. Since the degree of participating intention motivated by trust is higher for posters than for viewers, the website management should direct their resources more on posters with respect to security so that their trust in virtual communities remains strong. Reciprocity is insignificant for posters possibly because they do not feel the need of receiving help from others at this site. The web operator can consider establishing a direct connection between content providers and receivers. For example, an area can be designated where users can post certain questions on topics of interest to be answered by other users.

5.2 Limitations and future research

Several limitations of this study indicate the need for further studies. First, an extensive survey on multiple web services should provide more convincing results. Second, this study assumes that viewers and posters are two different types of user affected by social and personal factors. However, it is still unclear why and how such a transformation takes place. It could be caused by external factors such as social norms or stress, or certain internal characteristics that a person is more willing to share. Although the demographic data in this study shows no significant difference between viewers and posters, one study indicated that they differed in several aspects [6]. Further research is needed to understand whether or not the fundamental differences between posters and viewers exist or whether users change their status based on some other factors.

6 Acknowledgement

The author gratefully acknowledges the support of the Ministry of Science and Technology (Grant 102-2410-H-004 -198 -).

7 References

[1] Cooke, M., Buckley, N. (2008). Web 2.0, social networks and the future of market research, International Journal of Market Research, 50(2), 267-292.

[2] Hagel III, J., Armstrong, A. (1997). Net Gain: Expanding Markets Through Virtual Communities, Boston: Harvard Business School Press, MA.

[3] Williams, R.L., Cothrel, J. (2000). Four smart ways to run online communities, Sloan Management Review, 41(4), 81-91.

[4] Koh, J., Kim, Y.-G., Butter, B., Bock, G.-W. (2007). Encouraging participation in virtual communities, Communication of the ACM, 50(2), 69-73.

[5] Kollock, P., Smith, M. (1996). Managing the virtual commons: Cooperation and conflict in computer communities. In S. Herring (Ed.), Computer-mediated Communication: Logistics, Social, and Cross-cultural Perspectives, John Benjamins, Amsterdam, 109-128.

[6] Nonnecke, B., Andrews, D., Preece, J. (2006). Non-public and public online community participation: Needs, attitudes and behavior, Electronic Commerce Research, 6(1), p. 7-20.

[7] Preece, J., Abras, C., Maloney-Krichmar, D. (2004). Designing and evaluating online communities: Research speaks to emerging practice. International Journal of Web-based Communities 1(1), 2-18.

[8] Preece, J., Nonnecke, B., Andrews, D. (2004). The top five reasons for lurking: Improving community experiences for everyone, Computers in Human Behavior, 20(2), 201-223.

[9] Wasko, M.M., Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice, MIS Quarterly, 29(1), 35-57.

[10] Nahapiet, J., Ghoshal, S. (1998). Social capital, and the organizational advantage, Academy of Management Review, 23(2), 242-266.

[11] Kankanhalli, Q., Tan, B.C.Y., Wei, K-K. (2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation, MIS Quarterly, 29(1), 113-143.

[12] Ridings, C., Gefen, D., Arinze, B. (2006). Psychological barriers: Lurker and poster motivation and behavior in online communities, Communications of AIS 18(1), 329-354.

[13] Cho, N., Li, G.Z., Su, C.-J. (2007). An empirical study on the effect of individual factors on knowledge sharing by knowledge type, Journal of Global Business & Technology, 3(2), 1-15.

[14] Constant, D., Kiesler, S., Sproull, L. (1994). What's mine is ours, or is it? A study of attitudes about information sharing, Information Systems Research, 5(4), 400-421.

[15] Putman, R.D. (1995). Turning in, turning out: The strange disappearance of social capital in American, Political Science and Politics, 28(4), 663-683.

[16] Tsai, W., Ghoshal, S. (1998). Social capital and value creation: An empirical study of intrafirm networks, Academy of Management Journal, 41(4), 464-476.

[17] Dholakia, U.M., Bagozzi, R.P., Pearo, L.K. (2004). A social influence model of consumer participation in network- and small-group-based virtual communities, International Journal of Research in Marketing, 21(3), 241-263.

[18] Enders, A., Hungenberg, H., Denker, H-P., Mauch, S. (2008). The long tail of social networking: Revenue models of social networking sites, European Management Journal, 26(4), 199-211.

[19] Bauer, H.H., Grether, M. (2005). Virtual communities: Its contribution to customer relationships by providing social capital, Journal of Relationship Marketing, 4(1/2), 91-108.

[20] Porter, C.E., Donthu, N. (2008). Cultivating trust and harvesting value in virtual communities, Management Science 54(1), 113-128.

[21] Fukuyama, F. (1995). Trust: The Social Virtues & the Creation of Prosperity, New York: Free Press.

[22] Casalo, L.V., Flavian, C., Guinaliu, M. (2008). Promoting consumer's participation in virtual brand communities: A new paradigm in branding strategy, Journal of Marketing Communications, 14(1), 19-36.

[23] Blau, P.M. (1964). Exchange and Power in Social Life, New York: Transaction Publishers.

[24] Hars, A., Qu, S. (2002). Working for free? Motivations for participating in open-source projects, International Journal of Electronic Commerce, 6(3), 25-39.

[25] Chiu, C.-M., Hsu, M.-H., Wang, E.T.G. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories, Decision Support Systems, 42(3), 1872-1888.

[26] Lo, J., Chen, I., Lee, S. (2010). Top 100 list of Taiwan web sites in 2010, Business Next, 3(190), 83-89.

[27] Zaichkowsky, J.L. (1985). Measuring the involvement construct, Journal of Consumer Research 12(4), 341-352.

[28] Bagozzi, R.P., Edwards, J.R. (1998). A general approach for representing constructs in organizational research, Organizational Research Methods, 1(1), 45-87.

[29] Lee, H., Lee, Y., Yoo, D. (2000). The determinants of perceived service quality and its relationship with satisfaction, Journal of Services Marketing, 14(3), 217-231.Lin, H.-F. (2008). Determinants of successful virtual communities: Contributions from system characteristics and social factors, Information & Management, 45(8), 522-527.

[30] Cohen, D., Prusak, L. (2001). In Good Company: How Social Capital Makes Organizations Work, Harvard Business School Press, Boston, MA,.

[31] Thacher, J. B., George, J.F. (2004). Commitment, trust, and social involvement: An exploratory study of antecedents to web shopper loyalty, Journal of Organizational Computing and Electronic Commerce, 14(4), 243-268.

[32] Koh, J., Kim, Y.-G. (2003). Sense of virtual community: A conceptual framework and empirical validation, International Journal of Electronic Commerce, 8(2), 75-93.

SESSION

POSTERS

Chair(s)

TBA

Embracing e-Learning as a Catalyst for Enhancing Program Viability, Marketability, and Curricular Innovation

Adel Abunawass and Edwin Rudolph
Department of Computer Science, University of West Georgia, Carrollton, Georgia, USA

Abstract – *This poster addresses the ways in which transitioning a face-to-face graduate program in applied computer science to a 100% online program, where online education via e-learning is fully utilized and embraced, can function as a catalyst for enhancing and strengthening the long-term viability of the program; the ability to more effectively market the program to a wider population of potential students; and also encourage innovation in curriculum and pedagogy.*

Keywords: Online education; e-learning; computer science education; graduate education

1 Background

The University of West Georgia, a state-supported regional comprehensive university, established a campus-based face-to-face Master of Science in Applied Computer Science in 2002. From its inception, this program was designed primarily for students without an undergraduate background in computer science desiring to "re-train" and "re-equip" for a career in the computing and information technology field. The curriculum focused on providing a basic foundation in computer science, coupled with applied knowledge and skills in selected areas, and further allowed the student to design a program of study tailored to their individual interests. Students could choose to obtain breadth in a variety of areas through a program consisting of all course work, or engage in significant independent projects, or complete a traditional masters thesis. A small, teaching-focused computer science faculty delivered the program. The faculty also has responsibility for delivering an undergraduate program that is accredited by the Computing Accreditation Commission of ABET.

The program attracted a small but steady number of students during the first several years of being offered. Most of the students were local, and virtually all students obtained employment in the field shortly after graduation.

In recent years, however, it became clear that the program in its current form was not sustainable. This conclusion was the result of evaluation of assessment data, which indicated that the program's existing curriculum structure made it difficult to effectively evaluate student learning; reaching a plateau in enrollment; and shrinking availability of resources, primarily faculty. This was at a time when the department was simultaneously experiencing significant growth in its undergraduate program.

2 Embracing e-Learning for Change

In order to deliver the program in a more effective and efficient manner with already limited resources, more effectively market to and attract a larger population of students, and strengthen the financial viability of the program, the following strategic changes were identified: transition the program to 100% online within 3 years; modify the curriculum of the program to focus on a targeted set of learning outcomes; and adopt an innovative tuition model that would be affordable and at the same time enhance the marketability and sustainability of the program by providing a direct source of self-supporting revenue that would be directly used to provide support services to students in the program.

2.1 Obtaining Faculty Support

In order to garner faculty support and to prepare for the transition to 100% online delivery, we began with a gradual approach. The first step, even before substantial planning for the transition had begun, was to introduce a special initiative that would provide summer support for faculty to fully embrace e-learning by re-building and updating existing courses for online delivery. This initiative was carried over a 3-year period, and resulted in 10 re-designed courses built specifically for 100% online delivery. Importantly, these courses were not merely copied into an electronic format (i.e., recording the same lectures that would be delivered in a face-to-face format and posting materials to a website); rather, they were built specifically to take advantage of best practices for distance education and e-learning methodologies and pedagogy, and to utilize current technologies available to enhance the e-learning experience (e.g. learning management system, instant messaging, distributed source code control systems). Furthermore, the courses were designed for portability and reusability so that virtually any faculty member in the department could deliver the course as well as revise and update it in the future as necessary to accommodate changes in the subject matter.

2.2 Focusing the Curriculum

To address previous concerns relating to difficulty in evaluating student learning, as well as to support our intent to enhance the marketability and expand the potential student population, we needed to revise the curriculum to provide a clear and focused set of student outcomes. The existing curriculum allowed for too much variation in individual students' program of study for us to effectively assess student learning, as well as market the program as being able to provide students with a focused and specific set of knowledge and skills upon graduation.

Based on available faculty expertise and the needs of the job market, we determined that preparing students to enter the workforce as an entry-level software developer would be the focus of the revised curriculum for the program. The new curriculum (consisting of courses already prepared for online delivery as a result of the aforementioned special initiative) focuses on programming and software development, coupled with essential ancillary knowledge areas: web technologies, database systems, and an introduction to system and network administration. Students gain experience working in groups and on teams through two significant project experiences.

Whereas in the previous curriculum students could complete courses in any order, the new curriculum enforces a prerequisite structure to ensure students complete courses in the appropriate order and in a timely manner. Under the previous curriculum, students were able to, and frequently did, linger in the program well beyond the intended two-year duration of the program. This created challenges for student learning and resulted in a negative impact on progression and graduation rates. Federal and state mandates are also focusing increasingly more on college completion. So, it was important we do as much as possible to encourage students to complete the program on time.

2.3 Affordability and Sustainability

In addition to revising the curriculum to better market the program, we also wanted to make the program affordable—especially to out-of-state students who would now be a more important population as the program transitioned to 100% online delivery.

To achieve this, we were able to take advantage of innovative and flexible tuition policies provided by our state university system designed to encourage and support distance education. These policies essentially allow students who enroll only in 100% online courses to pay in-state tuition rates regardless of their place of residence; furthermore, they also avoid several other mandatory fees. Additionally, these policies allow for a differential tuition rate to be added to the base in-state tuition, provided the combined base and differential rate does not exceed the out-of-state rate. This differential can be used to provide resources directly to the program. Our program decided to adopt a modest differential rate which we channel directly back into supporting the program and its students. These resources provide for equipment and teaching assistants who assist faculty and provide a robust tutoring and technical support function for students in the program. Notably, the teaching support function not only alleviates issues with limited faculty resources, but also enhances the marketability of the program by addressing the needs of students who may work full-time by providing a range of flexible weekend and "after-hours" support. Finally, this tuition revenue enhances the program's sustainability by making it more self-supporting.

2.4 Curricular and Pedagogical Innovations

Innovation in curriculum and pedagogy has been a notable by-product of this transition and has had direct impact on both our online graduate program as well as our face-to-face undergraduate program. The effort and experience put forth by faculty to design courses for students they will never meet face-to-face encourages careful and thoughtful design of course material; this has for example resulted in increased hybridity of face-to-face undergraduate courses, enabling students to review and practice material outside of class meetings and enabling better use of scarce class time. Faculty who have developed "mini-lectures" (short, targeted videos on a particular topic) for online graduate courses are increasingly adopting this practice to supplement material and lectures in their undergraduate courses. It has also resulted in portability and reusability of content between undergraduate and graduate courses; for example, the department offers introductory system and network administration courses in both programs. Although the undergraduate course is not online, a large amount of the material and support tools developed for the graduate offering has been reused in the undergraduate.

3 Conclusion

Our experience has been that fully embracing e-learning and online program delivery, coupled with both institutional and faculty support, is an extremely effective catalyst for enhancing program viability as well encouraging innovation in both online as well as face-to-face and hybrid programs. Since moving the graduate program online, we have more than quadrupled the number of applicants and more than doubled the number of students in the program.

We are excited about the future of our program, and for online education and e-learning generally. We are especially excited that our graduate program was top-ranked by US News & World Report as a "Best Online Graduate Computer and Information Technology" for the last two years. It is particularly gratifying that this national ranking and recognition, which are of notable value for helping to market and the program, would not have otherwise been possible without fully embracing online education and e-learning to transform the program.

SESSION

LATE BREAKING PAPERS AND POSITION PAPERS: E-LEARNING, E-BUSINESS, ENTERPRISE INFORMATION SYSTEMS, AND E-GOVERNMENT

Chair(s)

Prof. Hamid R. Arabnia

A real time monitoring system to measure the quality of the Italian Public Administration websites

F. Donini[1], **M. Martinelli**[1] **and L. Vasarelli**[1]
[1]Institute of Informatics and Telematics, National Research Council (CNR), Pisa, Italy

Abstract -*In 2013, the Institute of Informatics and Telematics of the National Research Council (IIT-CNR) and the Department for Digitalization and Technological Innovation of the Presidency of Council of Ministers, carried out a pilot project, whose main aim was to create a monitoring system in order to verify and analyze the capacity of Italian Public Administrations to activate and manage information and services by means of the Internet. This paper describes the SmartPA (System for the Real Time Monitoring and Analysis of the Italian Public Administrations) application, developed by IIT-CNR, which analyses and monitors Italian PAs, evaluating in real time conformity and coherence of contents with Guidelines regarding PA websites, issued in the years 2010 and 2013. SmartPA, through the publication and presentation of results on a public portal, aims to establish an open and transparent communication between the citizen and Public Administration, thus bringing about a process of continuous quality improvement of PA websites*.

Keywords: e-Government, websites quality, case studies, websites monitoring tools, e-Transparency

1 Background

Within the 2012 e-Government plan and in the context of modernization of Public Administration, the 2009/11/26 Directive of the Ministry for the Public Administration and Innovation [1] [2], in order to rationalize Public Administration websites (PA) and to improve the quality of services and on-line information for the citizen, sets out specific Guidelines. These are aimed at providing clear indications regarding general criteria and operating tools for the rationalization of online content, for the reduction of obsolete public websites and improvement of active sites [3].

These Guidelines, which are an essential feature of the initiatives aimed at innovating the PAs, have the objective of establishing an open and transparent communication with the citizen, so bringing about a process of continuous quality improvement of public websites.

In particular, the Guidelines pay special attention to the definition of a map of minimum essential requisites. These features, on the basis of current legislation (e.g. Digital Administration Code, CiVIT Guidelines, Garante della Privacy, etc.), must be included within institutional PA websites.

Within this context, this project foresees the creation of a pilot monitoring system (SmartPA - System for the Real Time Monitoring and Analysis of the Italian Public Administration websites). This system, starting from *.it* domain names registered by Italian PAs[1], enables verification and analysis of the capacity to activate/manage information and services by means of the Internet and web channels. This evaluation has been carried out through the study and definition of a concise indicator of PA websites, measured in terms of coherence or compliance with regards to a minimum set of compulsory prerequisite contents, indicated in the Guidelines. This includes, merely as an example, the sections: organization chart; external relations office (URP); transparency; evaluation and merit of personnel; administrative procedures; public tenders; job competitions; on-line services; legal notices; certificate electronic mail (PEC).

SmartPA, exploiting the potential offered by ICT technologies, therefore provides real time measuring and evaluation tools for the quality of public services. For some time this has been the focus of PA modernization, with a specific emphasis on closely involving citizens.

2 Architecture

The SmartPA architecture is of a modular type (Figure 1). The system is subdivided into two macro-components, the *engine* and the *portal*. The engine is the elaborating part of the system, whereas the portal is the interface that enables the management of the engine and consultation of the results elaborated.

The *engine* module, which uses the Akka framework [4] for the creation of concurrent elaborations [5] [6], implements all the SmartPA logic, collecting, elaborating and retrieving data. The website analysis has been made by means of the parsing library Beautiful Soup [7].

The *portal* module, created by means of the Play Framework [8] web application, thanks to the use of the most advanced

[1].it domain names are managed by the .it Registry, that is a service of IIT-CNR

web technologies (HTML5, CSS3, JavaScript, etc.), is compatible with the main Internet browsers and mobile devices.

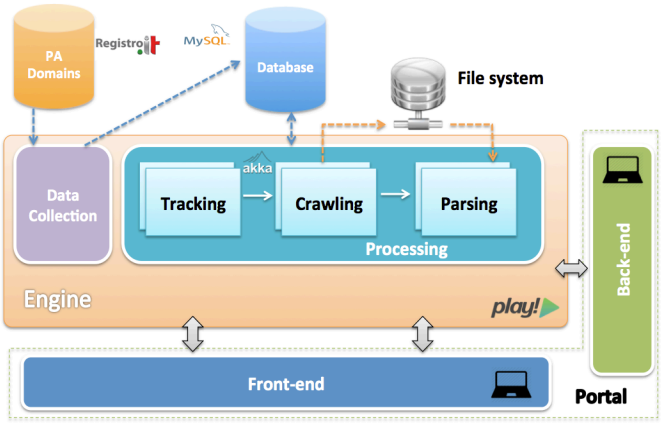


Figure 1 - SmartPA architecture

All the results (metadata, partial and total elaborations regarding analyzed subjects, references to downloaded web pages, etc.) are stored in a dedicated database. The html pages, which make up the monitored website, are stored in a file system, so as to enable future research and comparison. There now follows a detailed analysis of the two macro components that make up SmartPA: the *Engine* and the *Portal.*

3 The Engine

The SmartPA engine is subdivided into three main stages (Figure 2):

- *Data Collection* – this identifies the data set of relevant domain names (domain list) on the basis of which the analysis and monitoring of domain names is initiated;
- *Data Processing* – this uses the data of the Data Collection stage and is the key stage of the system. In fact, during this stage all the main activities and elaborations that make up the website analysis and monitoring process are carried out;
- *Data Discovery* – this is based on the Data Processing stage and identifies and associates a Public Administration with its specific institutional website.

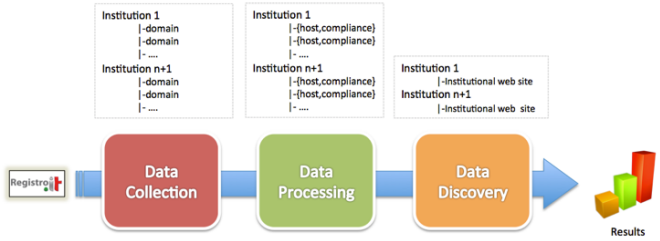


Figure 2 - SmartPA engine schema

The activity of Data Collection is mainly carried out during the initial stage of database population and in order to add new subjects to the monitoring process. On the other hand, Data Processing and Data Discovery activities can be initiated upon request of the user (real time) or periodically by the system (pre-established time intervals).

3.1 The Data Collection stage

Domain names, by their very nature, can be assigned to different types of registrants (public bodies, natural persons, companies, etc.) identifiable, in the .it Registry, by means of a specific univocal code. Using this code it was possible to retrieve the domain names assigned to Italian Public Administrations [9], to identify the institution they belong to, and for each registrant, identify the category (local institutions, research institutes, school, companies, local health authorities, etc.).

This activity also involved a preliminary process of data cleanup that gets Registry database data prepared to work correctly: in fact, various corrections were made regarding errors that had occurred during the registration of the domain names.

3.2 The Data Processing stage

On the basis of the data set of domain names obtained in the Data Collection stage, websites of each public administration were analyzed and conformity to the Guidelines were assessed. This phase involved a set of elaborations carried out by three ad hoc modules: the *tracker*, the *crawler* and the *parser* modules.

The *tracker* module identifies the URL of the homepage of the website associated with the domain and is able to filter replications of websites hosted by different domains, but belonging to the same PA. Particular attention was paid to the study of false positives such as white pages, splash screens, sites under construction, etc.

The *crawler* module uses as input the URL of the homepage obtained by the tracker and downloads the associated website. Given that the minimum contents foreseen by the Guidelines can be found also in second level subsections, for each website a download of at least three levels was carried out.

The *parser* module has the task of verifying the presence and compliancy of the contents foreseen by the Guidelines. For each content item the parser gives three different results (compliant, partially compliant, not compliant), in this way identifying a final score for each website.

3.3 The Data Discovery stage

Each PA can have multiple websites, but only one of these should correspond to the so-called *Institutional website,* the one that identifies the PA institutionally and highlights the structure, organization, staff, salaries, services offered, etc.

The purpose of the system at this stage is to identify the most likely institutional website of each PA. This relies on a scoring system (ranking) that takes into account, as weighted average, the results obtained by the parser (number of compliance to the Guidelines), the domain name pattern and the website URL pattern (Table 1).

Rank	URL	Comp.	Reason
1	http://www.comune.pisa.it	66	Institutional website (identified)
2	http://comune.pisa.it	66	Missing "www"
3	http://www.comune.pi.it	66	Not the best pattern "pi" less clear than "pisa"
4	http://cittapisa.it	3	Clearly not institutional for the numbers of compliances

Table 1 - Identification of the Institutional website according to the scoring system

4 Portal

The Portal consists of two modules and two separate sections, the public section dedicated to the presentation of data and statistical results (Front-end) and the system administration section (Back-end). The portal was designed to be intuitive, easily usable for the general public, accessible and equipped with a clear response layout for mobile devices (Figure 3).

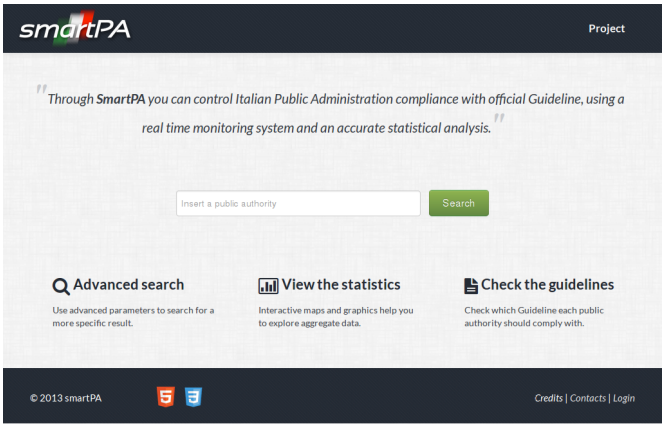


Figure 3 – SmartPA portal homepage

Access to the portal is therefore in two circumstances: in order to take advantage of the contents and results made available (as an anonymous visitor without credentials or as a user registered in the portal), or for its management (as administrator, operator and maintainer).

4.1 Front-end

The front-end module implements the portal section dedicated to the display of results. Among its main features this section includes the search (exact and partial) of a PA, and visualization by means of reports, charts and graphs of its descriptive information and compliance of the institutional website with the Guidelines (Figure 4).

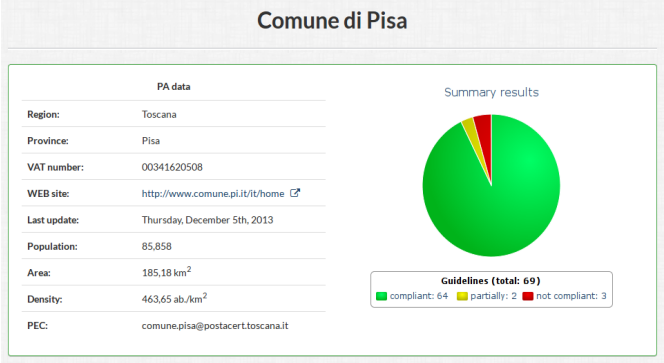


Figure 4 - Example of search result

This section also enables the user to search and, consequently display statistical aggregate data according to selectable parameters, such as the territory (on a national, regional and provincial basis), the type of PA or a specific Guideline (Figure 5).

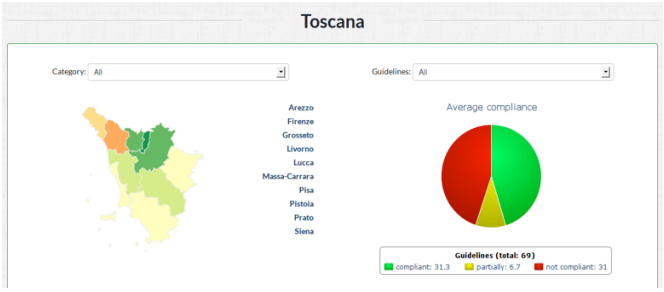


Figure 5- Results at a regional level

4.2 Back-end

The back-end module implements the section of the portal dedicated to system administration and interaction with the procedures of data collection. This section (Figure 6) offers an extensive set of features that enable the user to manage and monitor the main features of the system (information processing, domain names, organizations, users, system logs, general settings, etc.).

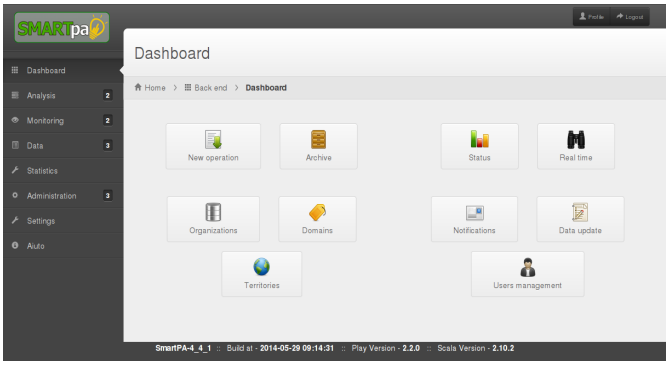


Figure 6 - Back-end dashboard

Through the back-end it is possible, for example, to launch Data Processing (3.2) and Data Discovery (3.3) processes and visualize the elaborations in real time, thanks to dedicated visual media. In the case of multiple instances of same time elaborations, they are managed by means of an ad hoc queue.

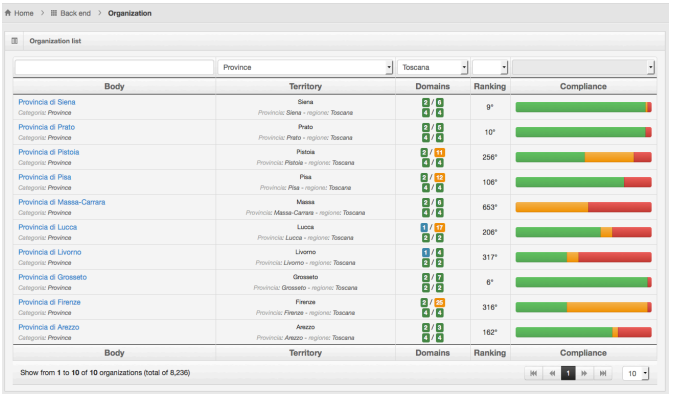


Figure 7 - Organization list detail

Other important features that are implemented in the back-end module are: visualization of the archive of the operations performed, consultation of the systems log, modifications of the profiles of the subjects undergoing verification (personal details, domain names associated with them, Institutional website URL, etc.) (Figure 7) and management of the system users.

5 Main results

The main statistical results obtained after the entry into operation of the analysis and monitoring system of the websites of the Italian PAs, are indicated below.

In January 2014 around 2,635,000 .it domain names were active, 34,115 (1.29%) of which were registered by Italian Public Administrations. 45% of these domain names (15,541) have been associated with subjects belonging to categories of interest of the present study (Councils, Provinces, Regions and Ministries). Within these subjects, over 8,232 PAs, 314 do not have a .it domain name. Some sample surveys showed that the main reasons for this situation are, for example, the availability of a domain name other than a .it (.es .eu, .net, .org), or the absence of a specific web site for the institution, etc.

Figure 8, shows the ratio between registered domain names and PublicAdministrationon a regional basis. Ideally, ratio should tend as much as possible to 1 as for the regions of Molise, Val d'Aosta, and Abruzzo. On the contrary, for what concerns regions as Tuscany, Emilia Romagna, Lazio, Marche the ratio turns out to be considerably higher.

Taking into account the analysis of the compliance of PA websites with the Guidelines, results show that on average, the degree of compliance is 47.7%. Therefore, around 33 Guidelines out of 69 are totally or partially respected in the

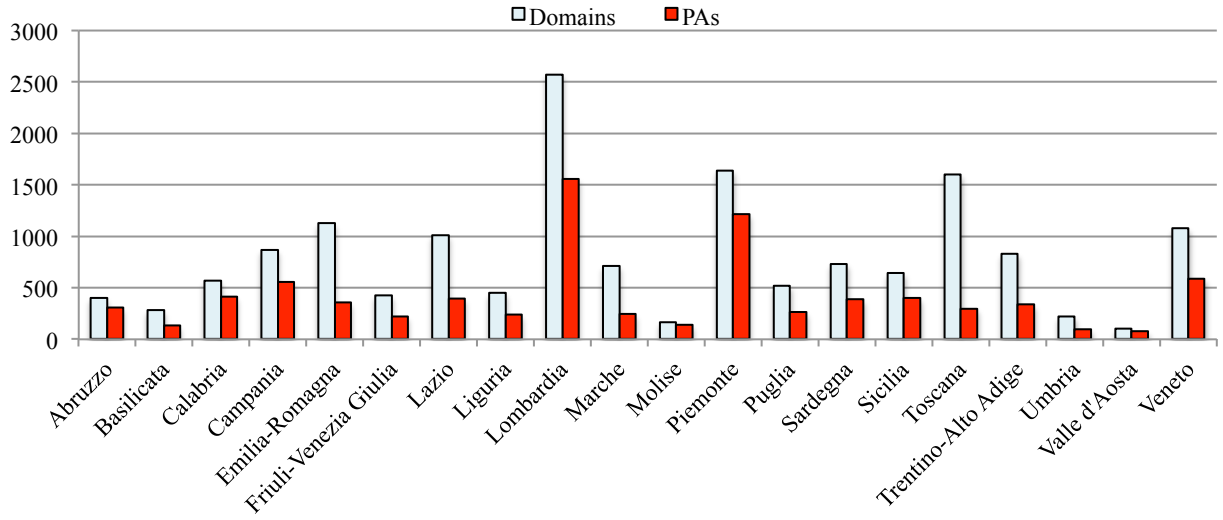


Figure 8 – PA distribution per region and domain names associated with them

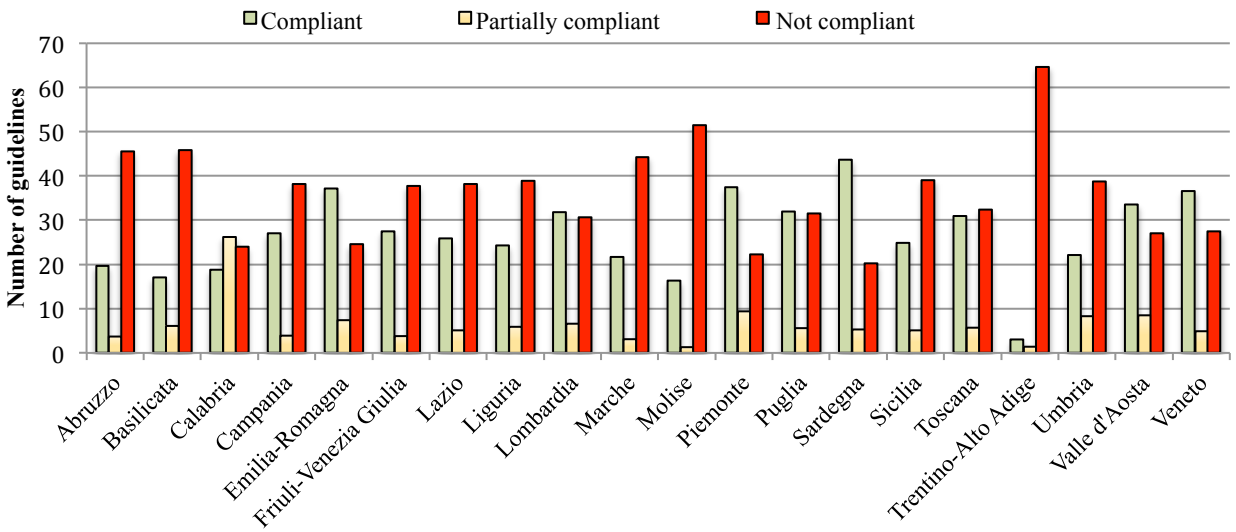


Figure 9 - Compliance with Guideline per region

websites of the PAs analyzed.

The highest average compliance at the regional level is achieved by the PAs of Sardegna (70.7%) region. On the other hand, the lowest average compliance can be found in Trentino Alto-Adige, with only 6.4% of compliant or partially compliant websites (Figure 9, Figure 11).

At a national level, results show (Figure 10) that most of the bodies have almost fully compliant websites or fully not compliant websites:

- 154 PAs have a website 100% compliant with the Guidelines (including 90 PAs that have a website partially compliant with the Guidelines);
- 1248 PAs have the institutional website that does not comply with any Guideline

Taking into account the situation at a Macro Area level (North, Centre and South), we can see that (Figure 11):

- PAs of the North have a degree of compliance with the Guidelines of 50.56% (35 guidelines out of 69 are respected);
- PAs of the Centre have a degree of compliance of 44.35% (31 out of 69 Guidelines are respected);
- PAs from the South and the Islands have a degree of compliance with the Guidelines of 48.19% (33 out of 69 Guidelines are respected).

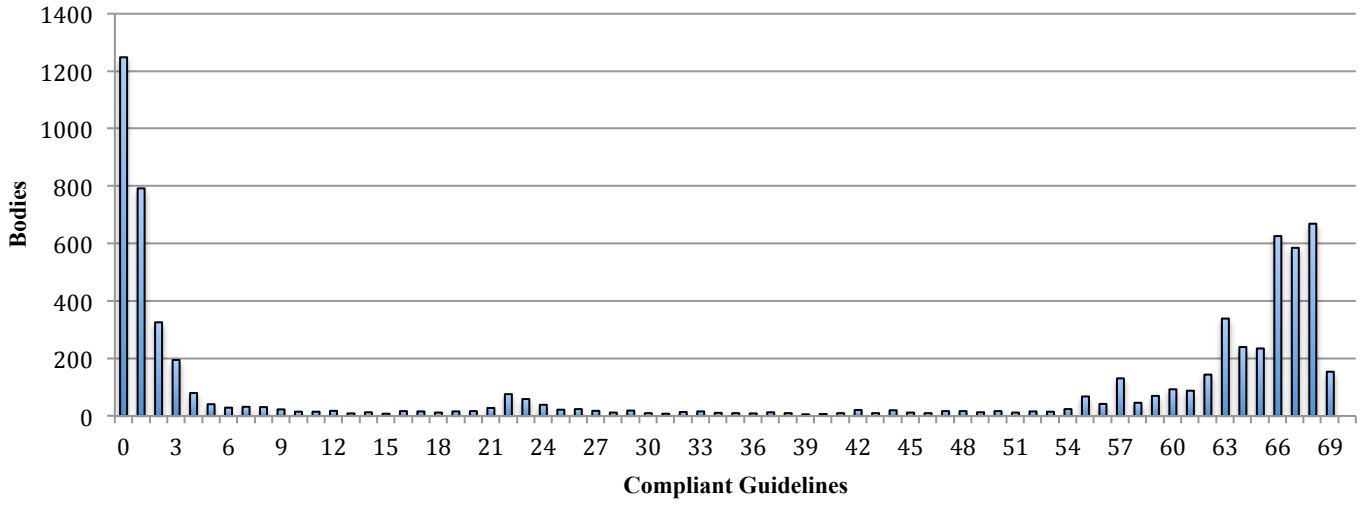


Figure 10 - Number of Bodies for Compliant Guidelines

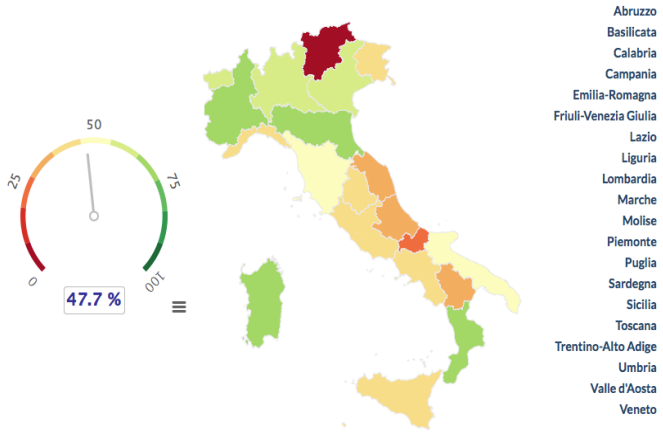


Figure 11- Compliance at national and regional level

6 Conclusions

The aim of SmartPA is to suggest criteria and tools for the rationalization of online contents, the reduction of obsolete public web sites and improvement of active sites.

Results show that in Italy there are still many Public Administrations that do not fulfill the requirements stated in the Legislative Decree dated 14 March 2013, n. 33.

This study also highlights the difficulties that a citizen may encounter while accessing a PA website, considering multiple domain names associated to it. In order to solve this problem, it is suggested the use of an electronic stamp that indicates the institutionalism of the reference web site of the body.

At a geographical level, there are no major differences for what concerns the quality of results. Only for Trentino Alto-Adige results differ from the national average; a possible explanation could be the presence of contents written in another language (German), considering the presence of bilingualism.

7 Future works

In order to contribute to the improvement of PA websites, SmartPA is constantly evolving.

A set of tools and advanced features are foreseen in order to increase awareness and, at the same time facilitate PAs during the phase of adjustment to this Decree. These features include an extension of the analysis and monitoring to all Italian PAs, an user-reserved dashboard that enable the PAs to manage their own profile and independently analyze the "quality" of their institutional website.

There will be also a tool for validating the accessibility of Institutional PA websites, able to verify adherence and compliance with the Guidelines for web content accessibility foreseen by the Stanca Law [10], the historical record of PA websites, able to show the trend over time of the quality of websites and services and the possibility to export of information and results in Open Data format [11].

Farther is foreseen a creation of a quality stamp to be assigned to the PAs with high levels of adherence and compliance with the Guidelines.

SmartPA is easily adaptable to other international e-Government contexts; due to its modularity, it allows the creation of custom rules to evaluate the quality of Public Administration websites.

8 References

[1] Directive n. 8/09 of the Minister of Public Administration and Innovation: http://www.funzionepubblica.gov.it/media/339253/dir_n_8_09.pdf

[2] Legislative Decree dated 14 March 2013, n. 33: http://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2013-04-05&atto.codiceRedazionale=13G00076

[3] Vademecum for website quality measurement: http://www.funzionepubblica.gov.it/media/1015979/vademecum-qualita-siti-web-acc.pdf

[4] Akka: http://akka.io

[5] Minish K. Gupta (2012). "*Akka Essentials*" – PacktPubblishing

[6] Venkat Subramaniam (2011). "*Programming Concurrency on the JVM – Mastering Synchronization, STM, and Actors*" – The pragmatic Bookshelf

[7] Beautiful Soup: www.crummy.com/software/BeautifulSoup

[8] Play Framework: http://www.playframework.com

[9] Codes of the councils, provinces and regions according to ISTAT : http://www.istat.it/it/archivio/6789

[10] Decree dated 20 March 2013, Amendments to Annex A of Decree dated 8 July 2005 : http://www.gazzettaufficiale.it/eli/id/2013/09/16/13A07492/sg

[11] Open Data, how to make open Public Administration data: http://www.dati.gov.it/sites/default/files/VademecumOpenData.pdf

[12] Fong, S. and Meng Ho Si, 2009. A web-based performance monitoring system for e-government services. *Proceedings of the 3rd international conference on Theory and practice of electronic governance*. Bogota, Colombia.

[13] Wang, L. and Bretschneider, S., 2005. Evaluating Web-based e-government services with a citizen-centric approach. *Proceedings of the 38th Hawaii International Conference on System Sciences*. IEEE.

Towards a REST-ful Visualization of Complex Event Streams and Patterns

Benjamin Kanagwa[1] **and Nasser Kimbugwe**[2]
Makerere University,P.O.BOx 7062, Kampala, Uganda
[1]bkanagwa@cis.mak.ac.ug, [2]nkimbugwe@cis.mak.ac.ug

Abstract— *This paper presents a scalable architecture for visualization of complex events, event hierarchies and relations. The goal is to support flexible, multi-level, multidimentional visualization of primitive and composite events to aid faster decision making. Our approach is to apply the concept of Representational State Transfer (REST) that focusses on design of large scale distributed systems. Different visualizations are first and foremost mapped to state representations of Complex Event Processing (CEP) systems. We show how RESTful design can be applied to common CEP visualization requirements such as root-cause analysis. Finally, we present a prototype implementation of the API using a case study of event streams in public procurement.*

Keywords: stream processig, complex event processing, visualization, Representational State Transter (REST).

1. Introduction

Diversity and responsiveness desired by users of complex event processing systems to explore large datasets in respect to different aspects has created a need for a more generic and scalable visualization infrastructure. We propose application of Representation State Transfer (REST)[12] to visualization in Complex Event Processing [16], [22]. REST is an architecture style designed for large scale distributed systems and has been applied in design of web services[8], [9], and other applications including twitter[1] and facebook[2] among others. The goal is to design an architecture that brings RESTâĂŹs advantages of loose coupling and good scalability to the discipline of CEP. The resulting uniformed interface provided by REST creates new opportunites including potential to integrated CEP systems across domains including mashups of CEP based systems. It has been noted by [21] that CEP is not yet fully exploited and is still establishing itself within the business world due to its lack of scalability and interoperability both within the event detection technologies as well as the visualization techniques. Current CEP visualizations do not provide the flexibility to alter, adopt and analyze the data from different perspectives.

[1]http://twitter.com
[2]http://facebook.com

CEP based systems continuously process events as they happen to infer complex relations based on context and temporal relationships. An event is any important occurrence that is worth noting and represents the instance of an activity [17]. Each organization has many events that happen at different times, some related and others not. Occurrence of events in time-space creates a continuous *flow* of events leading to the notion of *event streams*. An event stream has been defined as an infinite sequence of events [1]. Therefore, organisations have many event streams flowing into different information systems. CEP does not look at individual events but seeks to detect presence or absence of given event combinations (complex events) tied togehter by temporal relations. The event sources may be internal of external to the organization. External sources may indlude social media [3], [4] among others. The role of a CEP system is to continuously listen to incomming event streams and filter out desired combinations of events on which aggregation functions are applied to infer trends and instant insights for decision making.

Current users of CEP systems interact through dashboards with generic graphical representations for all users. However, users require personalized visualizations and appropriate access controls where a single user may be interested in different representations of the same event. It has been noted that visualization of these events in real-time is not yet scalable. Moreover current approaches make it impossible to integrate different CEP systems. This is largely due different event, pattern and visual representations. We note that provision of information is one aspect that must be complimented with appropriate representations that facilitate faster decision making. Personalized visual tools should allow browsing and exploration of events and related causes. We contend that each user at whatever level has some decision to make. Accurate, and current information is therefore required by users in varying level of detail and sophistication.

This problem is further compounded by large organizations and cloud environments, where thousands of users may be interested in observing different patterns. This requires the system to provide scalable mechanisms for serving such large number of patterns. Distributed scalable systems have been addressed by the Representational State Transfer (REST) design – a resource oriented architecture designed for scalable distributed systems. Through its uniform in-

terface, application of REST[12] to CEP brings several advantages. First, it enables scalable management of growing number of resources in a consistent way. Second, it provides a simple and flexible way to build CEP driven integrations and applications through a uniform programmable interface.

The focus of this paper is to provide a scalable architecture for visualization of data streams. In this paper we make the following contributions (i) we identify and categorize important *resources* for visualization in CEP (ii) define an API for scalable visualization of CEP patterns (iii) present a prototype implementation of the API. The rest of the paper is organized as follows :- in Section 2 we give related work followed by background in Section 3. Section 4 is a discussion of the design goals followed by the structure of the REST API in Section 5. We present a prototype implementation in Section 6 followed by a discusion and conclusion in Section 7.

2. Related Work

Many CEP tools and platforms have emerged in recent years. This is attributed to its wide applicability [19] including domains of finance, health care, fraud detection, and online business among others. Most notable platforms are IBM System S[7], Oracle[18] and Esper[7]. The Esper language and processing algorithm are integrated into the Java and .Net (NEsper) as libraries. These platforms differ at different levels including (i) internal representation of events - where a variety of alternatives exits including extensible Markup Language(XML), JavaScript Object Notation (JSON), internal object representations that are native to specific programming languages such as Java (ii) internal processing technologies - where two dominant options are Finite State Automata (FSA) based techniques and Event processing Networks (EPN) (iii) the number of extent of visualization support provided.

Esper is one of the leading open source CEP engines where events are modelled as object instances that expose event properties through Java Bean-style getter methods. In Esper, event classes or interfaces do not have to be fully compliant to the Java Bean specification; however for the Esper engine to obtain event properties, the required Java Bean getter methods must be present. Event specification languages provide constructs to define relations between event streams. For instance, Esper has in built functions for pattern matching [11] ,which uses five operators that include the following:- *every* - Operators that control pattern finder creation and termination; *Logical operators* - or, and , not;*Temporal operators* - that operate on event order $\rightarrow$ (followed by); *guards and where* - conditions that filter out events and cause termination of the pattern finder. According to Esper [11] the Esper engine exceeds 500 000 event per second with engine latency below 3 microseconds

In [4], Samujjwal suggested an adaptive event stream processing (ESP) environment in which he explored the limitations of current ESP systems due to fixed pattern detection mechanism. However, his work was more on pattern detection than on the visualization of the events. Annett and Stroulia in [2] developed a REST application called Invenio that could geographically visualize aggregated music chart information. However, Invenio is not a CEP application.

3. Background

3.1 Representational State Transfer(REST)

Since the invention of the REST architecture by Roy[12], it has seen wide adoptability and practical implementations that include Twitter[6], Facebook[20], Google[13], [14] and Ruby on Rails [15] among others.The modern web is one instance of a RESTful-style architecture [12] although applications can include access to other styles of interactions. In most cases, the REST API is provided as means to allow developers who want to embed functionality into their applications. REST is an architectural style which is designed around the concept of a *resource*[12]. REST systems expose their data and functionality through resources identified by Uniform Resource Identifiers (URI). REST-style architectures consist of clients and servers. Clients make requests to servers and servers respond to their clients by acting upon each request and returning appropriate responses. Each request and response is considered as transfer of a representation of a resource - hence the name REpresentation State Transfer.

A resource is a conceptual mapping to a set of entities [12]. In practice, a resource is anything that is important enough to be referenced by itself. However, during interactions, resources are represented in an addressable format like HTML, XML, RSS , PDF and many others depending on the prevailing need. Interaction between clients and servers is constrained by *uniform interfaces* defined by a fixed set of verbs. For instance, Hyper Text Transfer Protocol (HTTP) based implementations use the standard HTTP methods: GET, HEAD, POST, PUT and DELETE. A uniform interface only offers a set of operations including ability to retrieve, change or create data among others. The simplicity of the uniform interface is important to REST, because it keeps the interaction between client and server as simple as possible.

Resources can be accessed via HTTP. For each resource, a unique identifier is provided and makes part of the Uniform Resource Locators (URL) for operations on the resource. With REST, developers are able to create âĂIJmashupâĂİ applications that aggregate numerous sources of information and promote rich user interaction [2]. As an example let us consider the Twitter API with the URL structure of `http://twitter.com/statuses/user_timeline.extension`.

3.2 Complex Event Processing

An event can be categorized as primitive or composite event [24]. A primitive event is directly observed in the system while a composite event is a concatenation of primitive events using *event algebra*. Primitive events also known as *atomic* occur instantaneously.

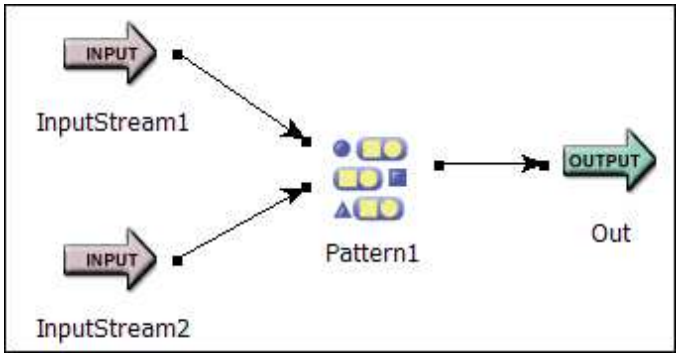


Fig. 1: Pattern matching(Adopted from [23])

CEP systems apply a query techninque called pattern matching where a set of event streams is matched against a complex pattern that specifies constraints on extent, order, values, and quantification of matching events [5]. This is mostly done using a pattern query which addresses a sequence of events that occur in order and are correlated based on the values of their attributes [1]. Figure 1 shows streams *InputSream1* and *InputStream2* that combined to detect pattern *Pattern1*. Upon detection of *Patttern1* an output *Out* is generated. Event stream queries provide the windows, aggregation, joining and analysis functions for use with streams of events[11]. As of today there is no standard pattern query language and each CEP engine uses either graphical interface, a scriptiing language or a combination of both. Pattern languages specify the constraints on the events to be detected and remain sialent on how the resultant event instances are to be visualized.

4. Requirements and Design Goals

Our design is based on a set of design goals as listed below:-

*R*1 - *Scalability*: CEP systems may be deployed at a small entity or very large entity. Therefore the number of concurrent users may vary from tens to several thousands. Rendering of graphs is time and computational intensive. It is therefore required that a CEP visualization infrastructure be designed with scalability.

*R*2 - *Abstraction*: The ability to fold and unfold, zoom in and zoom out different pieces of data is very important. This entails showing finer details of a piece of data when requested. Also abstraction allows hierarchical representation of events for different management levels. Aggregation and causality are fundamental concepts in CEP. Aggregation defines which set of primitive events combine to form a given complex event. The complex event is seen here as an abstraction of the primitive events. Causality involves sets of events that happen because another set of events had to happen. Both relations can be ably explored using abstraction techniques.

*R*3 - *Flexibility*: The ability to easily change the representation of data to provide quicker interpretations. This requires ability to transform from one visualization to another. For instance a user may want to render pie-chart or tabular visualization of the same data.

*R*4 - *Mashup*: Provide a basis for creating applications that aggregate different primitive and complex event patterns from disperate stream processing platforms

5. The REST API structure

The key consideration is that the appllication for which the the API is being designed for exists indpendently of the API. We identifty the states of the applications and operations that can be carried out.

5.1 Resources

As noted earlier, resources form the foundation of a RESTful design because interactions are based on transfer of resource representaions. One approach used to identify states and resources is to look at resources as equivalent to entities in an Entity Relation (ER) diagram. Following this approach, we identify the resources indicated in figure 2. These base resources are then augmented with derived resources such as lists of events, lists of activities, causal sets (history) that are perculiar to CEP visualization.

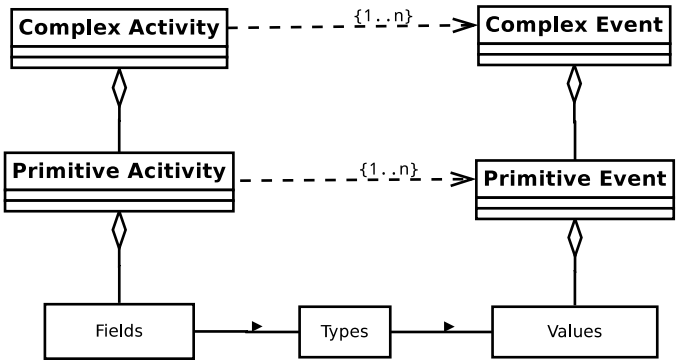


Fig. 2: Relationship between Complex Event Processing Resources

- *Activity* : These are event templates that define the *types* of events expected in the CEP system. The activity provide the event type and specifies the attributes and features of the event.
- *Event*: These occurrences representing that a specific activity has happened

- *Pattern*: group of activities that are constrained to happen under certain conditions based on ordering, extent, values and quantification of matching set.
- *Causal Models*: The events that cause an event to occur are bound to it [24]. They provide the information about the event through the parameters of the participating events.
- *Abstraction Hierarchies*: These represent how activities aggregate for form high-level activities
- *History*: a partially ordered set of event instances that contributed to the occurance of a given complex event.

5.2 URI

To illustrate the API, let E_1, E_2 be primitive event types. If events E_1 is required to happen before E_2 to match complex event E_5 then E_1, E_2 are components of complex event E_5.

5.2.1 Activity API

The general format for the activity API is `http://{activities}/{id}/format={fmt}` where `fmt` is one of `bar, piechart, linegraph, tabular`. This is further argumented with the HTTP methods as indicated in the table 1. The id represents the identity of the activity for which event instances belong.

HTTP Method	URI	Description
GET	/activities/	Returns a list of activities in the system.
GET	/activities/id	Returns a specific activity with corresponding id.
POST	/activities/id	Posts a given activity
PUT	/activities/id	Updates the activity with the specified id
DELETE	/activities/id	Deletes the event with the specified Id
GET	`/activities/id/components`	Returns a specific activity with corresponding id.

Table 1: REST API for Event Resources. Retrieves both complex and primitive activities

5.2.2 Event API

The general format for the events API is `http://{events}/{id}/format={fmt}`. This is further argumented with the HTTP methods as indicated in the table 2. The `id` represents the identity of the event.

Table 2 shows the events API. The last row is used to represent to retrieve the event history - the set of all events that led to its ocurrance. Consider the events of type E_1 that trigger events of type E_7 and E_7 in turn triggers E_9 , represented as $E_5 \mapsto E_7 \mapsto E_9$. When a match for E_9 is obtained then instances of E_7 and E_5 make up the history.

HTTP Method	URI	Description
GET	/events/	Returns a list of events the system.
GET	/events/id	Returns a specific event with corresponding id.
POST	/events/id	
PUT	/events/id	Updates the event with the specified id
DELETE	/events/id	Deletes the event with the specified Id
GET	`events/id/history`	the results would be e_1, e_2

Table 2: REST API for Event Resources. Retrieves both complex and primitive events

5.2.3 Zooming, Root cause analysis

The general format for the zooming API is `http://{zoom}/{id}/format={fmt}`. In this case the `id` represents either an acitivty or event. Complex events that result from pattern matching are also assigned a unique`id` and therefore can be accessed through the API. THe zooming API is applicable to both an activity or complex event. Zooming into a complex activity retrieves details that related to the hierarchical structure of the activity.

6. Implementation

We implementad our own engine called KAMANDA that is based on Non-deterministic Finate Automaton (NFA)[10]. The KAMANDA interface represents patterns using a graphical interfaces as indicated in Figure 3(a). The event sources are captures by adapters that connect to databases or some instrumentation. The engine is implementated using Java where events are modeled as JSON objects. The visualization interface is web-based in which we implement the RESTFul API. Events are made to pass through a CEP engine which filters them depending on the conditions set. All events, either directly from the CEP engine or from Database, are visualized through a REST visualization API accessible through the web browser.

Unlike many stream processing engines that use a combination of scripting and graphic interfaces for pattern language, the KAMANDA platform uses only a graphical user interface. This choice is based on the need to place control of information in the hands of managers and executives who are normally not IT experts and therefore find scripting languages hard to deal with. The graphical user interface provides several defaults that can be adjusted to provide more expressive pattern specitication options.

For illustration, we consider event streams E_1 and E_2 that represents `publish` and `sale` for an online auctioning system. The `publish` event represents an activity of a user publishing an item for sale, while the event `sale` represents an activity of sale of item to highest bidder. Figure 3 (a) shows how the pattern is represented. The conditions on

patterns are set through the expandable activity tabs. Figure 3 (b) shows the corresponding representations using a pie chart. In the same figure, a user can quickly choose between different representations. The segments in the pie chart is the total number of matching event patterns grouped every minute. So a number of 5 in the pie chart implies 5 pairs of `{publish,sale}` combinations.

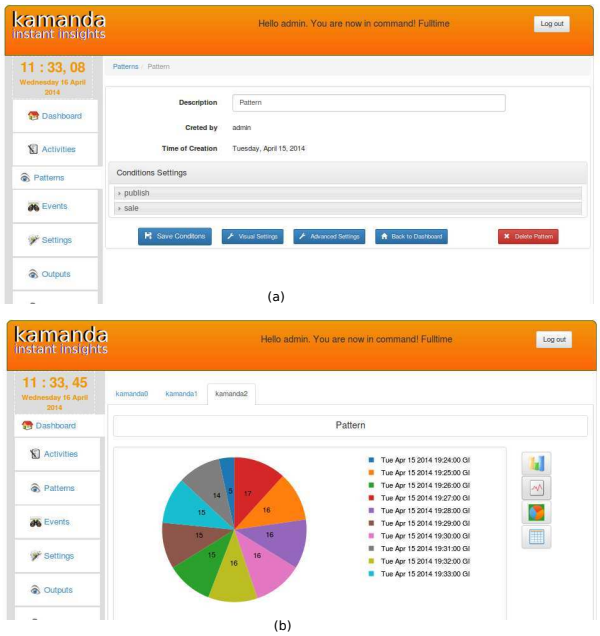


Fig. 3: API Implementation Pattern Detection

We provide two levels of root cause analysis through zooming. The first zoom level explores the details of aggregated value by window. The second level takes the invdividual results to show the specific event instances that make the original aggregated value. From a graph segment say in piechart one can zoom to see futher details which we show in Figure 4 (a). This is the first level of zooming. The second level of zooming will show the individual events as indicated in Figure 4 (b).

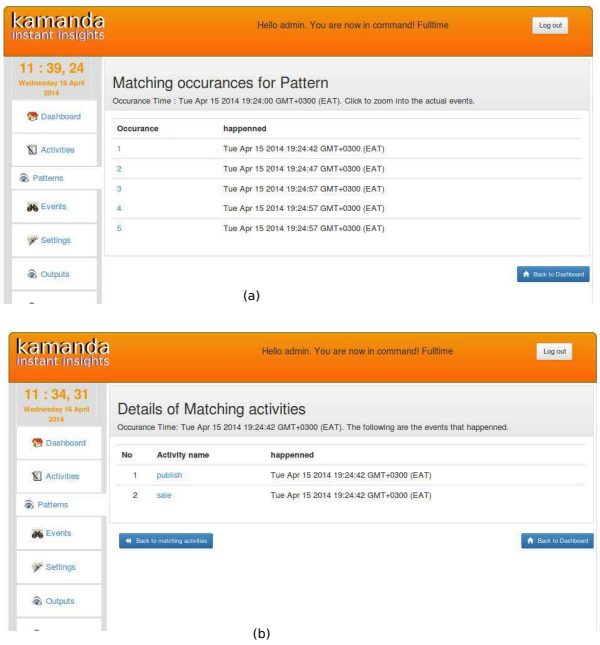


Fig. 4: API Implementation Zooming

7. Conclusion

We have been able to identify different resources akin to complex event processing. The resources have been exposed through a RESTful API allowing visualization. Through the proposed RESTful API and implementation, we have been able to provide a consistent and flexible means of navigating event stream outputs. Third party developers can embed CEP outputs int there systems. Our next step is to increase the expressiveness of the graphical pattern language as well the number of visual representations.

References

[1] Agrawal, J., Diao, Y., Gyllstrom, D., and Immerman, N. Efficient pattern matching over event streams. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (New York, NY, USA, 2008), SIGMOD '08, ACM, pp. 147–160.

[2] Annett, M., and Stroulia, E. Building highly-interactive, data-intensive, rest applications: the invenio experience. In *Proceedings of the 2008 conference of the center for advanced studies on collaborative research: meeting of minds* (New York, NY, USA, 2008), CASCON '08, ACM, pp. 15:192–15:206.

[3] Artikis, A., Etzion, O., Feldman, Z., and Fournier, F. Tutorial:event processing under uncertainty. In *Proceedings of the 6th ACM International Conference on Distributed Event-Based Systems* (New York, NY, USA, 2012), DEBS '12, ACM, pp. 32–43.

[4] Bhandari, S. An adaptive event stream processing environment. In *Proceedings of the on SIGMOD/PODS 2012 PhD Symposium* (New York, NY, USA, 2012), PhD '12, ACM, pp. 57–62.

[5] Cadonna, B., Gamper, J., and Böhlen, M. H. Sequenced event set pattern matching. In *Proceedings of the 14th International Conference on Extending Database Technology* (New York, NY, USA, 2011), EDBT/ICDT '11, ACM, pp. 33–44.

[6] Carey, B. M. Using the twitter rest api. *Explore the Twitter REST API for automated Web 2.0 downloaded from http://www.ibm.com/developerworks/xml/library/x-twitterREST/* (Jun 2009).

[7] Cugola, G., and Margara, A. Processing flows of information: From data stream to complex event processing. *ACM Comput. Surv. 44*, 3 (jun 2012), 15:1–15:62.

[8] Curbera, F., Duftler, M. J., Khalaf, R., Nagy, W., Mukhi, N., and Weerawarana, S. Ieee internet computing: Spotlight - unraveling the web services web: An introduction to soap, wsdl, and uddi. *IEEE Distributed Systems Online 3*, 4 (2002).

[9] Curbera, F., Khalaf, R., Mukhi, N., Tai, S., and Weerawarana, S. The next step in web services. *Commun. ACM 46*, 10 (2003), 29–34.

[10] Demers, A., Gehrke, J., Hong, M., Riedewald, M., and White, W. A general algebra and implementation for monitoring event streams, 2005.

[11] ESPER. Team: Esper reference documentation. $http : //esper.codehaus.org/esper - 4.6.0/doc/reference/en - US/html/index.html\%$ / (2012), accessed on 26th–August 2012.

[12] FIELDING, R. T., AND TAYLOR, R. N. Principled design of the modern web architecture. *ACM Trans. Internet Techn. 2*, 2 (2002), 115–150.

[13] GOOGLE. Google translate api using rest. In `https://developers.google.com/translate/v2/using\_restaccessedon28th/Jan/2013` (November 2012).

[14] GOOGLE. Using rest to invoke the api. In $https : //developers.google.com/custom - search/v1/using\_restaccessedon28th/Jan/2013$ (November 2012).

[15] HANSSON, D. H., ET AL. Ruby on rails. *Website. Projektseite: http://www. rubyonrails. org* (2009).

[16] LUCKHAM, D. The power of events: An introduction to complex event processing in distributed enterprise systems. *Ubiquity* (mar 2003).

[17] LUCKHAM, D. C. *The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.

[18] ORACLE. Event processing. *accessed from* $http : //www.oracle.com/technetwork/middleware/complex - event - processing/overview/index.html$ *on* (29th August 2012).

[19] PIETZUCH, P. Building secure event processing applications. In *Proceedings of the First International Workshop on Algorithms and Models for Distributed Event Processing* (New York, NY, USA, 2011), AlMoDEP '11, ACM, pp. 11–11.

[20] PURDY, D. Deprecating the rest api. In $https : //developers.facebook.com/blog/post/616/$ *accessed on 28th/Jan/2013*.

[21] SCHILLING, B., KOLDEHOFE, B., PLETAT, U., AND ROTHERMEL, K. Distributed heterogeneous event processing: enhancing scalability and interoperability of cep in an industrial context. In *Proceedings of the Fourth ACM International Conference on Distributed Event-Based Systems* (New York, NY, USA, 2010), DEBS '10, ACM, pp. 150–159.

[22] SEN, S., STOJANOVIC, N., AND SHEMRANI, B. F. Echopat: a system for real-time complex event pattern monitoring. In *Proceedings of the Fourth ACM International Conference on Distributed Event-Based Systems* (New York, NY, USA, 2010), DEBS '10, ACM, pp. 107–108.

[23] TIBBETTS, R., YANG, S., MACNEILL, R., AND RYDZEWSKI, D. Streambase liveview: Push-based real-time analytics. *StreamBase Systems* (Jan 2012).

[24] ZIMMER, D., AND UNLAND, R. On the semantics of complex events in active database management systems. In *Data Engineering, 1999. Proceedings., 15th International Conference on* (1999), IEEE, pp. 392–399.

Mobile Government Implementation for Government Service Delivery in Developing Countries: A South Africa Context

O.S. Ogunleye[1] **and J.P. Van Belle**[2]
[1] Information System Department, University of Cape Town, Rondebosch, Cape Town, South Africa
[2] Information System Department, University of Cape Town, Rondebosch, Cape Town, South Africa

Abstract - *Mobile phones are becoming the most rapidly adopted technology in history and the most popular and widespread personal technology in the world. Also, the ubiquitous nature of mobile phones in developing countries, for example South Africa provides an opportunity to use of this platform to provide better service delivery to the citizens of the developing countries. This paper identifies major service delivery issues in South Africa. Various m-government systems that have been implemented in other countries to solve service delivery issues were identified. A survey was circulated to citizens and government stakeholders*

Keywords: e-government; m-government; Developing Countries; Service Delivery; South Africa.

1 Introduction

Government can be defined as the dynamic mixture of goals, structures and functions by which a community is ruled [1]. However for a government to achieve its goals, various efforts and opportunities are utilized either through electronic means or traditional paper means. Electronic government (e-government) refers to the delivery of government services through the use of information and communication technologies (ICTs). Mobile government (m-government) which is seen as the subset of e-government is the delivery of government services through wireless technologies, anytime, anywhere by adopting diversity of mobile devices.

The challenges of government services delivery as well as the opportunities of adopting and using wireless and mobile technologies represent pressures and drivers to which modern governments should respond. However, the ubiquitous nature of mobile technologies (particularly mobile phones) in the developing countries (for example South Africa) provides an opportunity to use of this platform to provide better government service to the citizens. E-government is a cost-effective solution that improves communication between government agencies and the citizens by providing access to information and services through using mobile devices. M-government has attracted more and more research interest and focus from national governments and universities [2].

The potentials for m-government in developing countries, however remains largely unexploited. Different human, organizational and technological factors, issues and problems affecting these countries, require focused studies and appropriate approaches. Mobile technology has been referred as an "enabler", but on the other hand it should also be regarded as a challenge and a peril in itself. The organizations, public and private which ignore the potential value and use of mobile technology may suffer crucial competitive disadvantages. Nevertheless, some m-government initiatives have flourished in developing countries too, e.g. India, Kenya, China Brazil etc. The experience in these countries shows that government in the developing countries can effectively and efficiently utilize and appropriate the benefits that mobile technologies offer, but m-government success involves the accommodation of some unique conditions, needs and obstacles. The adaptive challenges of m-government implementation go beyond technology as they also call for organizational structures and skills, new forms of leadership, transformation of public institutions [3].

Moving away from these assertions, the primary aim of this paper is to identify and analyse the primary issues, opportunities and challenges that m-government initiatives present for developing countries. The insights and results presented here are based on an empirical research conducted in South Africa as well as a case studies undertaken in 10 developing countries (Brazil, Chile, China, Columbia, Guatemala, India, Jamaica, Kenya, the Philippines, and South Africa) which have already explored implemented m-government initiatives. In these cases we observe different applications and opportunities for m-government such as tax administration (Jamaica, Guatemala, South Africa), better services to customers businesses and stakeholders in general (Brazil, India), and m-government for transparency and business efficiency (The Philippines, India, Chile).

Furthermore, we investigate the primary m-government needs in South Africa. A survey was constructed regarding this as well as factors that will influence the successful adoption and use of these services. The survey was circulated to the South African citizens where access to traditional internet is limited and as such mobile device is the de facto computer in their hands. We also conduct interview of the stakeholders in various government agencies. We got an

overall implementation perspective on the domain as well as what is needed about the needs of the citizen in the implementation.

The research contributes to the existing body of knowledge in that it identifies specific factors that influence the adoption as well as use of m-government system in the South African context. The factors that are identified and discussed relate to the identified m-government services providing government and the general stakeholders with an overview of what is required to implement the similar services.

2 Literature Review

2.1 Focus Shift in the Public Sector

The emergent of the internet, digital connectivity, the explosion of mobile technology and the use of m-commerce and m-business in the private sector are pressuring the public sector to rethink hierarchical, bureaucratic organizational models. Customers, citizens and businesses in the developing countries are faced with new innovative m-business and m-commerce models implemented on a daily basis by the private sector and made possible by mobile technologies and applications. The same model has be required by the citizens in the developing countries. Citizens are customers for governments due to the fact governments need to empower rather than to serve, focus on prevention rather than to cure, be mission oriented and customer focused and shift to team work and participation rather than to hierarchy.

In recent times, public sector are beginning to understand the potential opportunities offered by the by the mobile and e-business models to fit with the citizens' demands, to offer better services to citizens and to also increase efficiency by streamlining internal process. This means that mobile technology causes a "focus shift" which introduce the age of "network intelligence" by reinventing business, government and individuals. This is because the traditional bureaucratic is being replaced by competitive, knowledge based economy requirements, such as speed up in service delivery, network organization, innovative entrepreneurship, customer driven strategy, flexibility and vertical/horizontal integration [4]. This new focus shift of the government toward m-government strategy, coordinated network building, customer services and external collaboration [5]

2.2 Defining M-government

M-government is a paradigm that means different things to different people. To some, m-government can simply be defined as digital government information or a way of engaging in digital transaction with customers through mobile devices and technologies. To others, m-government consists of the creation of mobile applications where government information about political and government issues is presented.

The definitions and conceptualization of m-government limit the range of opportunities that it offers. And one of the many reasons why m-government initiatives have failed is related to limited definition and lack of total understanding of the concept of m-government, its processes as well as its functions.

M-government is a complex and multifaceted concept that requires a broad definition of and understanding, in order to be able to design and implement a successful m-government strategy. However, the fundamental element of the various m-government definitions is the use of mobile technologies and tools to reinvent the public sector by changing its internal and external way of doing things as well as the relationship with the customers (i.e. the citizens) and the business communities. The analysis of these definitions allows us to individualize the main issue and components that characterize an m-government framework such as:

- Change areas (internal, external, and relational)
- Users, customers, actors and their relationships (citizens, business)
- M-government application domains (mServices, mDemocracy, mAdministration)

2.3 Change Initiatives

The definitions and analysis of m-government provided above incorporate three important change areas. These are:

Internal change, which is the adoption of mobile technologies to improve the efficiency and effectiveness of internal functions as well as processes of government by interrelating different departments and agencies in delivering service to the citizens. Therefore the flow of information among various government agencies and departments occurs at a faster rate, thus reducing processing time and paperwork bottlenecks.

External change which refers to the possibility of using mobile technologies and devices by government to be more transparent to citizens and business, thus providing access to the wide range of information gathered by the government. M-government will also generate opportunities for partnership and collaboration among many government institutions. The lines between within government agencies will not only be blurred but also the lines between the government and those that touch it.

Relational change which refers to the use of m-government implementation to enable the fundamental changes in the relationship between citizens and the government. Through this, vertical and horizontal integration of services and information from various government agencies and other stakeholders get seamless services.

These three change areas shows that m-government initiative does not consist of a simple business process reengineering. It requires a radical thinking of the nature and functioning of the organization and the relationships between

organizations which needs focus in a web of relationships including all levels of business functions in which the boundary inside and outside are permeable and fluid [6]

2.4 Relationship Network of M-Government

M-government aims to target four main groups. These are the citizens, governments (other governments and public agencies), employees and businesses. The transaction that occurs via the mobile platform between the government and each group constitute the m-government network of interactions.

The first group deals with the relationship between the government and the citizens which allows government agencies to communicate and relate with its citizens, which allows accountability, democracy and improvements to public services. The second group is the relationship between various governmental organizations, for example the national government, regional government and local government. This is due to the fact that governments depends on other levels of government within the nation in order to effectively deliver services and allocate responsibilities and in order to realize a single access point, collaboration and cooperation within different governments and agencies is compulsory.

The third group is the relationship that occurs between government and its employees which is an effective ways to bring various employees closer to themselves and the government in order to promote knowledge sharing among the employees. The final group constitutes the relationship between government agencies and private businesses. This allows transaction initiatives such as m-procurement through the use of mobile technologies.

The full exploitation and implementation of the relationship network in the implementation of the m-government initiatives requires the three applications domains of m-government.

2.5 M-government Applications Domain

The domain to which m-government implementation aims to target can be sub divided into three groups. These groups are:

- **mServices** which is aimed at realizing the connections and interrelationship between the government and the citizens in order to deliver automated government service through mobile devices.
- **mDemocracy** which helps to enable relationship and interaction between citizens and the government beyond boundaries
- **mAdministration** which is used to achieve the purpose of mobile automation and computerization of administrative tasks and for realization of strategic connections of internal processes and functions.

These applications domains are considered as overlapping and m-government can be found in the overlapping area of these three application domains thus demonstration the complexity and heterogeneities that are needed to be taken into consideration for ensuring its successful implementation. (Fig. 1)

2.6 M-government Initiative and Implementation in the Developing Countries

Informatics is a field of study that is primarily involved with the application of information, Information Systems and Information and Communication Technologies (ICTs) within organisations. Informatics can therefore be defined as the study of Information, Information Systems and Information and Communications Technologies which is pragmatic to various phenomena [7]. Following this definition of informatics, Government informatics can therefore be defined as the pragmatic use of information, information systems and information and communication technologies within Government organisations. This however incorporates application of m-Government - an extension of e-Government, which is primarily concerned with the delivery of Government services through mobile devices and mobile technologies [8].

Therefore it can be argued that the application of ICT causes a paradigm shift by introducing the age of network intelligence, reinventing businesses by Government and individuals [9]. In line with this, Governments around the world have taking steps towards implementing a wide range of ICT applications in the past decades. Countries have been classified by the United Nations according to the World Economic Situation Prospect (WESP) as developed economies, economies in transition and developing economies [10].

Developed economies include, for example, the United States, Canada, West European countries and Japan; economies in transitions include for example Croatia, Montenegro, Belarus while developing economies include for example Argentina, Brazil, India, Mexico, South Africa and Bulgaria, to mention just a few. For all countries, the application of ICTs for Government reinvention is increasing not only in investment but also in terms of increase in the number of high-profile initiatives that have been launched which are visible in the country. However, majority of these inventions have taken place in the developed countries. These countries are influenced by the fact that an information society will result in economic and social benefits [11]. According to [11], information infrastructures are projected to incite economic growth, increase productivity, create jobs, and improve on the quality of life. Furthermore, there is a big difference between ICT implementation and use between developed and developing countries.

However, there is lack of adequate and sufficient infrastructures such as computer, access to Internet, access to funds, etc. in developing countries. Therefore, the developed countries have an easier way of implementing ICT projects

such as E-Government than the developing countries. Mobile technologies, especially, mobile phones that can be used to access the internet and perform such activities beyond voice data alone have become the computers in the pocket of many citizens in the developing countries [12]. Due to this, concepts such as mobile Government implementation have gained priority, and thereby eliminate access restrictions. In the light of this, Governments in developing countries are trying to foster their capacities to be agile and ubiquitous. Therefore, they are slowly evolving service delivery towards mobile devices. However, this reality needs careful analysis, prototyping and evaluation of services to investigate whether any change leading to this forms of service delivery, and/or access, will be accepted by citizens and implemented by the Government.

2.7 M-government Readiness in South Africa

Public service delivery is an important and topical issue in any country. Citizens depend on the Government to deliver services effectively and efficiently. A true reflection of democracy has been displayed through the South African Government for more than a decade and many citizens in many developing countries have been restless because of election promises and manifestoes that have not been kept. In recent years, e-Government implementation has been highlighted as a significant contributor to public service delivery.

However this contribution is not directly visible to the majority of the public for which the services are meant for and much of e-Government take place "behind the scenes" and these "behind the scenes" involves activities which include storage of data and records pertaining to every aspect of citizens from birth to death. The South African national e-Government strategy addresses each of these phases in a citizen's life by developing e-Government services relate to each life phase [13].

Although, some South African provincial Governments, for example [13] have made significant progress towards the implementation of e-Government, the state of e-Government in South Africa at the national level however is still at the rudimentary stages [14]. There are various reasons for such an evaluation, the major reason being the lack of facilitation to update and adopt e-Government services by the majority of the citizens and an evaluation of the expectation of the citizens who are the primary users of the system [15].

Furthermore, indications are rife that most of the grumbled communities will not be appeased through the implementation of e-Government. This is because, this will not always be evident and directly visible to the citizens largely due to lack of internet connectivity [15]. However, Machiavelli [16] argue that "*There is nothing more difficult to take in hand, more perilous to conduct, or more uncertain in its success, than to take the lead in the introduction of a new order of things*". This argument also applies to the adoption of mobile and wireless technologies in the delivery of Government service to the citizens.

More than 93% of South Africans have access to mobile phones while 90% are owners of mobile phones [17]. This high penetration of mobile technologies presents an opportunity to reach an exceptionally broad base of citizens in the developing countries (for example South Africa). This has raised motivation for mobile Government service implementation.

Also, Government has the responsibility to deliver quality service and information to the citizens at all levels of life. These services and information that are sometimes critical are needed by citizens in making decisions and forming opinion. This helps them to feel a part of the Government [18]. This also allows timely service delivery to the public and therefore helps to promote public participation in democracy and creates accountability and transparency [18]. To this end, mobile technology has proven to be a critical channel through which the Government delivers services and information to the citizen. This is called Government to citizen (G2C) service delivery [18] and citizen communication with the Government (C2G and M-democracy) [19]. This also includes Government's delivery of service to the business (G2B) and business interacting with the Government (B2G).

3 Research Methodology

The aim of this research is to identify and analyse the primary issues, opportunities and challenges that m-government initiatives present for developing countries. After the empirical research was conducted which was based on the case studies undertaken in other developing countries, we turned to the research question: ***How can the Government incorporate the use of mobile cellular technologies to improve the provision and reach of government services?***

An interpretivist philosophy was adopted in this research. The research instrument that was selected for the purpose of this research is a questionnaire survey which consists of both qualitative and quantitative questions. This is available from the authors upon request. The decision to arrive at the questionnaire that was developed is based on the Besides the demographic information questions, the instrument has a section gauging on how the citizens would like to use their mobile devices to interact with the government if given the opportunity in South Africa. A final section uses the constructs taken from the UTAUT model to gauge what factors might influence the success of mobile government implementation and how important are these constructs. One of the key factor to the adoption of m-government implementation by the citizens was identified as: at what cost a citizen would be using his /her mobile devices for m-government system? Furthermore, we conduct an interview of the stakeholders, majority of who are senior government employees in order to gauge information on the

factors that may be responsible success and failure of m-government implementation in the developing countries.

The majority of the population identified for data collection include citizens from the bottom of the pyramid who do not have access to traditional internet connectivity and who are living in the rural communities but have access to mobile phone that can be used to access the internet. The sample consist at least two rural communities in 7 provinces of South Africa (South Africa has 9 provinces in total). These respondents have been selected as suitable due to the fact that they constitute the largest number of those who need access to one or more government services and as such our research objectives directly point to them. There was a particular focus on citizens between ages 18 – 35.

3.1 Data Analysis

3.1.1 Respondent Demographics

Responses were collected from more than 750 citizens during the data collection process and 653 of these were included in the dataset. The other responses were excluded due to one error or the other. Furthermore data was collected from 5 senior government employees.

3.1.2 Respondent Analysis

Respondents were asked various questions in relations to how they would love to interact with the government in terms of the government service delivery. After the elicitation of demographic information, we gathered information about the capability of the mobile phones of the responded. We gathered that 74 percent of the respondent have mobile phones that has capabilities of either smart phones or feature phones. On the issue of the convenient of using their phones to communicate with the government. 81 percent of the total respondents explained that they feel convenient using their phones in this manner. With participation in politics, 77 percent said that they will be willing to participate in politics with their mobile phones.

Furthermore, 90 percent of the total respondents said that they feel secured using their phones to communicate or transact with the government. Further informal interviews on this reveals that respondents feel that, their information is submitted to the government one way or the other during transaction with the government. On the issue of the reason why respondents would prefer using their mobile phones to communicate with the government. The majority of the respondents (95 percent) agrees that it saves times, money and Queue as it is convenient. The researcher further enquire about barrier that will not allow them to adopt m-government system implementation. From this, we gathered that language and cost will be barrier to m-government system implementation. We further investigated this. Therefore, we review some academic literature regarding the successful adoption of mobile government system, we find out that the major bone of contention, truly is the question of who pays for what. However, 68 percent respondents said that they are willing to bear the cost as long as it will save them time and energy.

4 Conclusion

The infrastructures and realities of Africa in general and South Africa in specific present various challenges for the integration of Mobile-Government in the way government runs its business to ensure quality service delivery and better dividends of democracy. Mobile phones are largely a pervasive presence in the larger part of the community and it although not free, is stable and nearly ubiquitous. There are numerous advantages for using existing technology and infrastructure for government to ensure quality service delivery. However, little research has been done on the use of mobile technology in social research in an African context which is differ distinctively from the rest of the world as the technology is not primarily used for the mobility that it offers but for the ability to access communities, collaborate and communicate. There are very limited applications on mobile available to support government business in ensuring quality service deliveries. This research study has highlighted some of the potential ways in which m-government systems implementation will help empowerment the citizens in the developing countries.

Therefore, m-government implementation will create an opportunity for African and other developing countries to take advantage of mobile phones as a crucial tool for empowerment and development in developing countries, Africa in particular. It will enable a standardized and framework-based approach to delivering government services to citizens, thereby enhancing citizens' participation in the democratic process. This can only be achieved by utilizing various functionalities of mobile phones, in ways that make sense in Africa, while the issues addressed above are taking into proper consideration. Although, this may seem to be a very African approach, we are convinced that the use of the mobile phone as the computing tool of choice in government service delivery will quickly be emulated in other developing countries. This is due to the fact that these devices are becoming more powerful, more ubiquitous and even more multi-functional.

5 References

[1] T. A. Pardo. (2000, March 19). *Realizing the Promise of Digital Government: It's More than Building a Web Site.* Available: http://demo.ctg.albany.edu/publications/journals/realizing_the_promise/realizing_the_promise.pdf

[2] L. Carter and F. Belanger, "The utilization of e-government services: Citizen trust, innovation and acceptance factors.," *Information Systems Journal, 15,* vol. 15, pp. 5-25, 2005.

[3] A. B. Allen, L. Juillet, G. Paquet, and J. Roy, "E-Governance and Government Online in Canada:

Partnerships, People and prospects," *Government Information Quaterly*, vol. 18, pp. 93–104, 2001.

[4] O. S. Ogunleye and J. P. Van Belle, "Exploring the Success, Failure and Factors Influencing M-Government Implementation in Developing Countries," in *IST-Africa*, Mauritius, 2014.

[5] T. K. Ho, "Reinventing Local Governments and the E-Government Initiative," *Public Administration Review*, vol. 62, pp. 434-444, 2002.

[6] D. Tapscott, *The Digital Economy*. New York: McGraw Hill, 1996.

[7] P. Beynon-Davies, *Information systems : An introduction to informatics in organisations*. Basingstoke: Palgrave, 2002.

[8] F. Bannister and D. Remenyi, "The Societal Value of ICT: First Steps Towards an Evaluation Framework," *Journal of Information Systems Evaluation*, vol. 6, pp. 97-206, 2003.

[9] V. D. Ndou, "E-Government in Developing Countries: Opportuninties and Challenges," *Electronic Journal on Information Systems in Developing Countries*, vol. 18, pp. 1-24, 2004.

[10] UnitedNations. (2012, March 19). *Statistical Annex: Country Classification.* . Available: http://www.un.org/en/development/desa/policy/wesp/wesp_current/2012country_class.pdf

[11] L. V. Audenhove, "Information and communication technology policy in Africa: A critical analysis of rhetoric and practice," in *Information Technology in Context: Studies from the Perspective of Developing Countries*, C. Avgerou and G. Walsham, Eds., ed Burlington, USA: Ashgate Publishing company, 2000,, 2000.

[12] O. S. Ogunleye, "Context and Capability: The Future of Small Screen Research and Development in Africa," presented at the M-government Consortium International Conference, Brighton, UK, 2010.

[13] WesterncapeProvince. (2012, October 20). *E-government Strategy.* . Available: http://www.westerncape.gov.za/text/2012/10/wcg-draft-e-government-strategy-for-public-comment-october-2012.pdf

[14] S. M. Mutula and J. Mostert, "e-Government in South Africa: e-service quality access and adoption factors," *Electronic Library*, vol. vol. 28, pp. 38-53, 2010, 2010.

[15] G. Kaisara and S. Pather, "e-Government in South Africa: e-service quality access and adoption factors," in *World Wide Web Applications*, Port Elizabeth, South Africa, 2009.

[16] B. M. Network. (2014). *Brainy Qoute*. Available: http://www.brainyquote.com/quotes/quotes/n/niccolomac131418.html

[17] C. Rawlinson. (2011, February 15). *Infographics: Cellphone Usage SA Stats.* Available: http://www.chrisrawlinson.com/2011/02/infographic-cellphone-usage-sa-mobile-stats/.

[18] A. G. Farshid and I. Kushchu. (2004). *M-Government: Cases of Developing Countries.* Available: http://www.m-government.org/resurces/mgovlab_afgik.pdf

[19] H. Brücher and P. Baumberger, "Using Mobile Technology to Support eDemocracy," presented at the International Conference on System Sciences, Hawaii, 2003.

Using Linked Open Data to Enrich a Corporate Memory of Universities

Cristal Karina Galindo Durán[1], R. Carolina Medina-Ramírez[2] and Mihaela Juganaru-Mathieu[3]
[1, 2] Departamento de Ingeniería Eléctrica, Universidad Autónoma Metropolitana, Unidad Iztapalapa, D.F. México.
[3] Institut H. Fayol, École Nationale Supérieure des Mines de Saint Étienne, Saint Étienne, France.

Abstract - *In this work we describe a proposal for managing educational resources produced by three high level educational institutions. Our approach takes into account not only the meta informations inside educational resources but also the meta informations extracted from linked open data in the web associated to massive open online courses (MOOCs). We visualize educational resources as a semantic corporate memory where semantic technologies and Linked Open Data approaches can improve the information integration. Text mining takes an important part in our approach for generating links between concepts and detecting some important associations.*

Keywords: Link Open Data, corporate memory, MOOC, text mining, Semantic web

1 Introduction

A corporate memory (CM) is an explicit and persistent representation of knowledge inside an organization, which can be embodied in persons or documents both with heterogeneous content and formats (doc, xls, ppt, etc.) [1]. In particular, these documents may contain data from different areas of knowledge, so document management requires in a broad sense an intelligent storage and management in order to drawn on the knowledge inside them and links with large pools of domain data, personal or social data. We consider as Semantic Resource Integration (SIR) the search process and significant data retrieval existing inside information resources (documents, people) available in a digital format. This integration is based on the use of semantic technologies and linked open data approaches in order to retrieve pertinent documents or navigate between linked documents. Our approach considers an educational corporate memory (ECM) case study: the semantic integration of educational resources produced by three high level educational institutions: École Nationale Supérieure des Mines de Saint Étienne (FR), Universidad Autónoma Metropolitana, Unidad Iztapalapa (Mx) and Escuela Militar de Ingenieros (Mx), see the Figure 1 to look the main functions of this proposal.

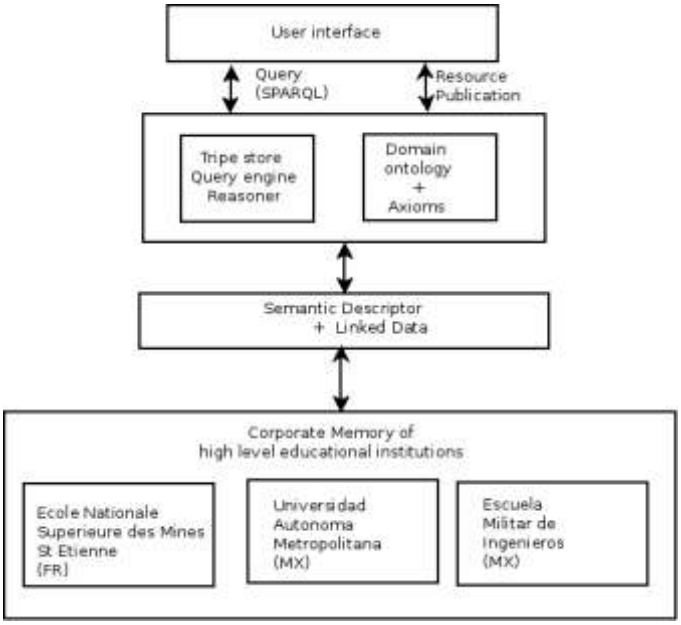


Figure 1. The main components of our proposal system able to manage a corporate memory and to integrate Link Open Data.

We also take into account: the explosive growth of educational resources, heterogeneity in format, content and structure, concepts ambiguity, among others. The main users of the ECM are students, lecturers, professors and people who are related to the educational domain. These educational resources can be consulted in order to retrieve specific documents or people according to a concept, or something that is well known, finding information that is not explicit in educational resources.

Although document characterization and document access can be simple to implement in the case of a centralized or distributed corporate cloud system, information search is not immediate, some text mining functionalities are needed in order to automate the educational resource characterization process: for example classification to detect the main domain, the level and the language of a MOOC or other external document, named entity recognition to detect the keywords, mining links to detect the area and the level of an educational resource. On the other hand, this organizational memory can be enriched with documents and external information to organizations from other web sources such as: funding

bodies, enterprises, customers, press documents, documents from partners or descriptive semantic information.

The ability to easily access the full corporate memory and the presence of semantic search indexes and resources are conditions that greatly simplify the entire process of mining. In Figure 2, we show a representation of a variety educational resources and indexes associated. Automatic generation of educational resource descriptions helps the automatic creation of indexes.

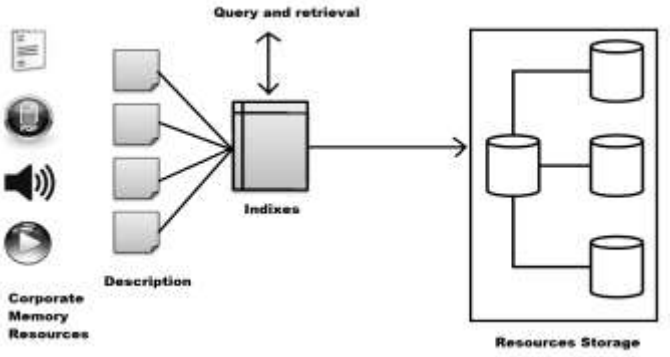


Figure 2. Indexes representing the retrieval information

Linked Open Data has a vital role in open education where MOOCs are the main protagonists. The open education is emerging as a collective web of linked educational data about courses, resources, platforms, instructional models, experts, professors, and so on [2]. However, all these initiatives do not converge yet to a cohesive and truly connected global view [2]. Universities are joining the movement by deploying their own Linked Data platforms, and many other organizations and initiatives are also exploring how reuse, integrate and interoperate isolated Open Educational Resources (OER) repositories using Linked Data [3]. Our proposal tries to contribute to the Open Educational Resource integration in order to be a truly connected global view of OER repositories and MOOCs. Metadata enables a level of interoperability between different learning platforms [4], there are currently no standards or metadata exchange in *massive open online course***s** (MOOCs)**,** we consider it as a corporate memory and thus to enable resource sharing across different platforms. Indexing is a crucial phase in Open Educational Resources retrieval, so text mining techniques could help to automate this phase. After obtaining indices we can apply semantic technologies and linked open data, which have shown the feasibility of formal representation of knowledge in a specific domain for seeking information considering the content of the resource. The use of ontologies, axioms, reasoner and semantic descriptions of educational corporate memory resources are significantly increase the quality of response of information search systems and open educational resource integration.

The rest of the paper is structured as follows: Section II introduces some concepts that will be used throughout the document. Section III describes an overview of our proposal a complete system for managing an Open Educational Corporate Memory using Semantic Web technologies and Linked Open Data (LODCMS). In section IV, we describe the process to be followed for obtaining educational resource descriptions and then propose semantic index which could guide the educational resources storage and retrieval. Finally, in Section IV some conclusions and future work are presented.

2 Theoretical and explanatory framework

The MOOCs are distinguished from other online courses, because they provide academic support as well as support and guidance to students throughout the course. The main features are: a) Mass: We do not known exactly how many students will be and may be more than can be met in person; b) Open: materials are available for free; c) Online: Profit of different communications channels and tools offered by Internet.

There are two main streams on MOOCs [5] considering educational and organizational models, these categories are: cMooC and xMOOC. cMOOC are developed and conducted by academics through open courses and web platforms. They are based on a participatory and collaborative approach. xMOOC: is a kind of of MOOC where the teacher is still the center of instruction, she/he prepares lectures, makes online discussions and others activities.

Many initiatives have been emerged around MOOCs and produced some platforms such as: Coursera [6], Udacity [7] EDx [8] MiriadaX [9], among others.

Linked Data refers to data published and reusable on the Web and readable by a machine. Its meaning defines that are linked to external data sets [10] [11]. Open Data refers to the legal interoperability of data [12]. Keep in mind that the use of linked data approach does not require the use of open data, however in order to get a potential profit, it is necessary to publish data as Linked Open Data good practices.

The open data format for publication is the Resource Description Framework (RDF) [13]. In 2006 Berners-Lee established four rules for publishing data on the Web; these rules are known as "Open Data Principles" [14] and consist of:

1. Use URIs to identify resources published on the Web.
2. Use HTTP URIs and so people can find those names.
3. When someone looks for a URI, provide useful information, using the standards (RDF, SPARQL, etc.).
4. Include links to other URIs, so that they can discover more things.

The process for looking for a Web URI in order to obtain information on a referenced resource is called dereferenced [15]. URIs can identify resources which could be [16]: a) a resource that can be transmitted through the Web, b) concrete or abstract entities that cannot be transmitted through the Web.

Once Open Data is published, we can perform various operations such as: storage, visualization, query,

linking data, reasoning on data to name a few. Figure 3 shows graphically the operations on Linked Open Data.

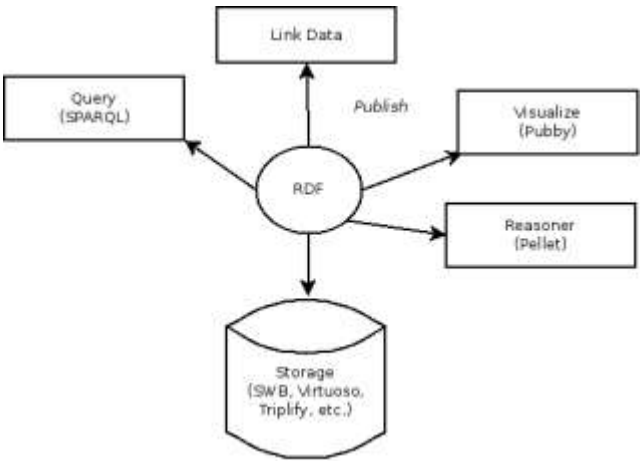


Figure 3. Operations on Linked Open Data

3 Overview LODCMS

The interest of this work is focus on the knowledge and resources in three educational organizations (teaching and research), in which there are complete knowledge about their members, internal documents, etc., and incomplete information from external sources (e.g. funding bodies, corporate web sites). The problem is studied from the perspective of data integration (expressed in RDF standard), where we have a domain ontology (representing in OWL format) as a central knowledge model for validating data. We can use reasoning on domain ontology and data (instances) to infer relationships between resources in the memory.

The data mining techniques [17] or text [18] help directly for classifying or clustering, supervised and unsupervised techniques respectively. Also other text mining techniques will be very useful to find general patterns or features in collections or groups of documents and serve to confirm or discover new semantic links, for this reason the links mining techniques [19] could be applied as well.

All mining methods (text, data or links) are based on document content (structured or unstructured). It is essential to generate indexes related to educational resources available in the corporate memory considering the availability, integrity and privacy. The use of a corporate cloud seems a good approach to storage and manage this indexes. Others approaches like platform (PaaS model) [20] or type service (SaaS) [21] are also considered.

Our goal is to propose a complete system for managing an Educational Corporate Memory using Semantic Web technologies (LODCMS). In particular, open linked data and ontologies. Running LODCMS imposes the existence of modules and the analysis of the dialogue with the user, collections management, generation and management of resources corresponding to the indices corporate memory and motor resolution of questions which in turn interacts with engines: information retrieval, reasoning and data mining. Another module would be able to decide if the result delivered to the user (the answer) is worth or not worth to be kept permanently. This decision will depend on the feed-back from the same user.

In Figure 4 we show user interactions with LODCMS and other interactions with the outside. The user will ask the system and have an answer (like a document, a list of documents, a collection, a link or an ontology entry, and also the user could send her/his feed-back and indicate, if the response is correct or meaningful. The administrator naturally manages the whole system and is able to integrate new semantical resource. The system can automatically detect the new internal documents and ontologies and also the changes of the extra-resources that are known and can do its self-alimentation.

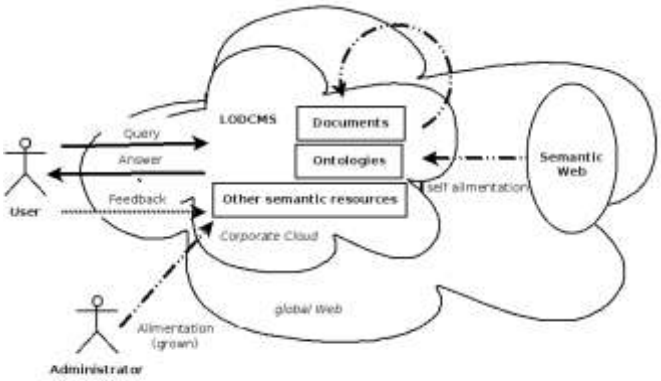


Figure 4.- The interaction with the outside of LODCMS: users and Semantic Web.

The resources contain semantic ontologies, thesauri, dictionaries, wikis, layout databases (relational or not). The data contained in the documents refer to internal resources (own institution) or outside (other web resources). Ontologies are much more complete and general resources, and are easy to handle, for this reason in the system there is in a gap between ontologies and other semantic resources that are more heterogeneous.

The auto-feed functionality would treat external data (global web) and internal data (semantic resources themselves, information resources, and ontologies) in a unified way. Indexing latter to access the first one, through links, reaching the globality of interesting information (global web). The indices denote therefore necessary to access the data (documents and ontologies) as links refer to a very fine way to a part of the collections or documents and the ability indices equally treat all data.

In figure 5 we indicate the whole data mining process which contains the mainly phases of a classic mining process: preprocess, choosing parameters, methods, interpretation and evaluation of the results, feedback, and also our mining process contains a feedback grown of the semantic resources we used. The phases as preprocessing, more exactly

representing a textual document or a structured document, can be better done using semantic resources as the wikis and thesaurus or bag of named entity.

Not only one step of mining is done, but also successive steps in order to detect some important information like language, domain (as biology, mathematics, etc.), areas of domain (as trigonometry, cellular biology, etc.), topics and levels of domain (as college, high education). Automatically, if language and domain are missing, a classification step can be done to detect this information. To detect areas and subareas of a domain some various mining process can be done as named entity recognition, classification inside a well-known resource, topic modeling.

In the mining process we can also detect some patterns after a topic modeling process, a clustering between a small collection and an association rules algorithm or detect some interesting link using link analysis (see [22] for more details), this new knowledge after an expert validation step can be integrate inside our indexes and semantic resources.

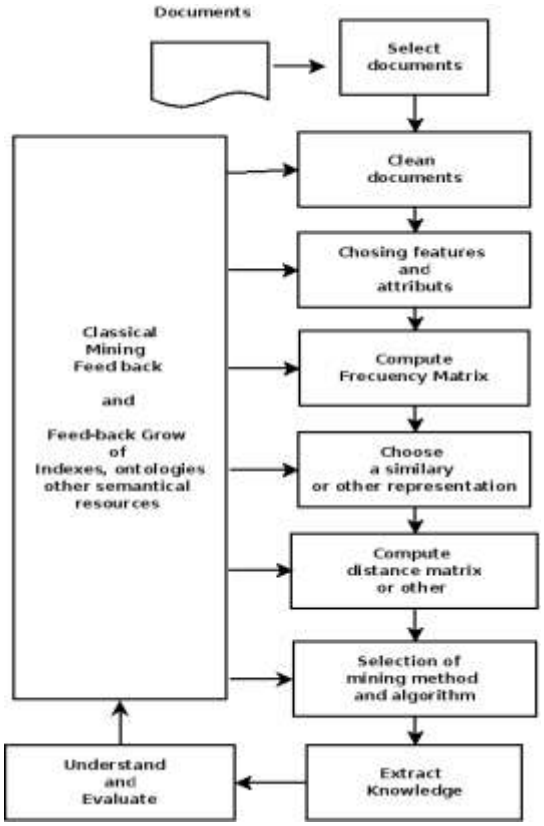


Figure 5. Proposed methodology for the whole mining process by generating also indexes and links between documents and terms.

4 Conclusion and future work

A proposal for semantic resource management of an organizational memory (LODCMS) is described in this article. The contribution of semantic technologies and linked open data was emphasized. We can also mention that our LODCMS system may be able to generate data that would be opened and linked to others, both to increase the visibility of the organization, to support the growing development of the Semantic Web.

A methodology for obtaining descriptions and index generation, based on the concept of clustering is presented, and the importance of these for the inclusion of semantic Web technologies showed.

5 References

[1] Alarcón Zamora, Erik "Integración Semántica de Recursos de Información En Una Memoria Corporativa". Master Degree Thesis in Science and Technologies of Information, UAM-Iztapalapa., Mexico January 2014.

[2] Mathieu d'Aquin and Stefan Dietze "Open Education: A Growing, High Impact Area for Linked Open Data", *ERCIM*, No 96, pages 14-15, http://ercim-news.ercim.eu/en96/special/open-education-a-growing-high-impact-area-for-linked-open-data

[3] Nelson Piedra, Janneth Alexandra Chicaiza & Jorge López "An Architecture based on Linked Data technologies for the Integration and reuse of OER in MOOCs Context", *Open Praxis*, vol. 6 issue 2, April–June 2014, pp. 171–187

[4] Aguirre,S., Quemada, J. and Salvachua, J. "Mediadores e Interoperabilidad en E-learning". In *V Conferencia Internacional Anual sobre Educación, Capacitación Profesional y Tecnologías de la Información Virtual. Educa 2004. Forum Universal de las Culturas.* Barcelona,Spain.

[5] Martín, C. M. V. "¿Serán los COMA (MOOC), el futuro del e-learning y el punto de inflexión del sistema educativo actual?" SCOPEO, (2). URL: http://scopeo.usal.es/seran-los-coma-mooc-el-futuro-del-e-learning-y-el-punto-de-inflexion-del-sistema-educativo-actual/

[6] Coursera Websites, 2014. URL: https://www.coursera.org/.

[7] Udacity Websites, 2014. URL: www.udacity.com.

[8] EDx Websites, 2014. URL: https://www.edx.org/.

[9] MiriadaX Websites, 2014. URL: https://www.miriadax.net/.

[10] Berners-Lee, T., Hendler, J., and Lassila, O. "Linked Data-The Story So Far". *International Journal on Semantic Web and Information Systems* (IJSWIS), 5 (3):1-22. March 2009.

[11] Short Guide of Linked Open Data, 2014. URL: http://www.w3c.es/Divulgacion/GuiasBreves/LinkedData

[12] Lybrary Linked Data Incubator Group Final Report, 2011. URL: http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/

[13] Hjelm, J. "Creating the Semantic Web with a RDF". Wyley, 2001.

[14] Linked Data- Design Issues, 2013. URL: http://www.w3.org/DesignIssues/LinkedData.html

[15] How to Publish Linked Data on the Web, 2014. URL: http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/

[16] Labra-Gayo, J.E. "Web Semántica: Comprendiendo el camino hacia la Web 3.0 ". NetBiblio. 2012.

[17] Ramírez, M. J. y Ferri, C. "Introducción a la Minería de Datos", tomo 1. Pearson Prentice Hall, 1ª ed ., 2004.

Srivastava, A. N. y Mehran, S. "Text Mining Classification, Clustering and Applications". CRC Press Taylor & Francis Group, 2009

[18] Getoor, L. y Diehl, C. "Link mining: a survey". *SIGKDD Explor. Newsletters*. 7, 2, pags 3-12. December 2005 DOI=10.1145/1117454.1117456.

[19] Yi Wei; Blake, M.B., "Service-Oriented Computing and Cloud Computing: Challenges and Opportunities," *Internet Computing, IEEE* , vol.14, no.6, pp.72,75, Nov.-Dec. 2010. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5617062&isnumber=5617049

[20] Loutas, N.; Kamateri, E.; Tarabanis, K., "A Semantic Interoperability Framework for Cloud Platform as a Service," *Cloud Computing Technology and Science (CloudCom), 2011 IEEE Third International Conference on* , vol., no., pp.280,287, Nov. 29 2011-Dec. 1 2011 http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6133154&isnumber=6133119

[21] Ronen Feldman, James Sanger "*Text Mining. Advanced Approaches in Analyzing Unstructured Data*", Cambridge University Press, 2007.

Characteristics of the Operational Dimension for the Bid Process Information System

Sahbi Zahaf
Higher Institute of Computer and Multimedia
MIRACL Laboratory, Sfax University, Tunisia
sahbi@zahaf.net

Faiez Gargouri
Higher Institute of Computer and Multimedia
MIRACL Laboratory, Sfax University, Tunisia
faiez.gargouri@isimsf.rnu.tn

***Abstract*— Bid process translates the techno-economic expertise, which partners build in a cooperative way. It is a key business process which evaluates the results of different trade tasks: hence, it influences the company's survival and strategic orientations. Therefore, the Information System that supports this process must be characterized by integrity, flexibility and interoperability. Nevertheless, the urbanization approach, on which we rely to implement this system, has to deal with "three fit" problems. To overcome these problems, we suggest addressing these exigencies following an operational dimension which remains responsive to other dimensions: the organizational and decision-making ones. However, the cooperative dimension covers the remaining dimensions. It ensures the consistency and the interaction between the different dimensions. We are interested on the characteristics of operational dimension.**

Keywords- Bid process, Information System, ERP.

I. INTRODUCTION

The owner calls for a bid to benefit from a product or service; after some period, he receives many proposals from different participating companies which submit their responses to this call for tenders (bid process). A bid process embodies a techno-economic proposal (a technical expertise backed by a financial offer). Such a contribution translates preliminary, the recommendations proposed by each contributor, either to reconstruct the desired product, or to organize the required service. Therefore, it is an elementary study that takes place before negotiating the contract with the owner, i.e. before launching the project. The bidder (company that pilot the bid process) might appeal to some partners, especially during the construction of the technical proposal. In the meantime, we can have new calls for bid following the first one and so on. Each participant seeks to cover and meet mainly his financial benefits because it is recognized as the most powerful and incorporating dimension. It interrelates with and touches the other dimensions like the: technical, social, environmental, etc. In each evolutionary step of the bid process, collaborators can renounce their participation when they recognize that such a deal does not cover their objectives. Indeed, most of the participants who decide to pass their last proposal, make a thorough evaluation exercise that tests the feasibility and effectiveness of the offer, through the technical, economic, temporal, and risk indicators. The bid process is a key business process of the company insofar as it directly affects its future and locates its expertise and its competitiveness in relation to its competitor [1]. The owner focuses on the best offer, which meets its requirements and which covers the eminent interests. Once gained the offer, the bidder and his collaborators create the bid process team. The latter makes part of the industrial consortium which targets to realize the project in practice.

Certainly, it is the company's agility and competence that allows acquiring the customer's confidence and interest, and as a consequence wins the offer. It is to remember that doing business means taking risks, something that can influence the company's growth and survival. So, the strategic management of business is a major and current concern to an innovative enterprise. The latter promotes to restructure its Information System (I.S) around its trades and business processes.

The I.S is the executing support of the processes of business. It directly influences the internal and external environmental requirements of the company. Internally, the consistency of an I.S depends utterly on its degree of integrity, flexibility and its internal interoperability. While externally, I.S agility always depends on its flexible capacities and external interoperability. Nevertheless, the reality shows agility and consistency problems on both inter as well as intra levels. We prove this premise when we apply the urbanization approach [9]. It consists of three transposition problems dubbed as "*three fit*": (i) "*vertical fit*" (lack of integrity and lack of extensibility); (ii) "*horizontal fit*" (lack of flexibility and lack of internal interoperability); (iii) "*transversal fit*" (lack of openness and lack of external interoperability).

Our objective is to suggest I.S which supports the bid process (B.P.I.S.). This system must be: *integrated*, *flexible* and *interoperable*. We treat these aspects following four dimensions: (i) the *operational dimension* which aim to specify the bid operation process designated for a specific project; (ii) the *organizational dimension* which targets to specify the set of skills and knowledge that the company had acquired due to previous bids in order to eventually reutilize this patrimony in future bid projects; (iii) the *decision-making dimension* whose goal is to optimize decisions on the market's supplies; and finally, (iv) the *cooperative dimension* which seeks to plan the inter-enterprises cooperation that take place during the construction of the techno-economic proposal of the offer. In this work, we are interested on the operational dimension.

This article is organized as follows. First, we propose our modelling of bid process. Second, we identify the required characteristics I.S supporting a bid process, and we focus on the hindrances encountered while implementing I.S: problems of "*three fit*". Then, we suggest an urbanized I.S (B.P.I.S.), to overcome these limitations. After that, we specify the characteristics of the operational dimension of our B.P.I.S. We finish this work by a conclusion that opens the horizons to our future research projects.

II. BID PROCESS MODELLING

This section deals with bid process modeling using BPMN according to the PRIMA approach [1], which distinguishes three main steps when modeling bid processes: assessment of the bid eligibility, elaboration of the bid process, and closure of the bid process. Thus, bid process exploitation is achieved through the interaction of these three sub-processes (Fig. 1).

Once the company receives a bid, it first checks the eligibility of the bid. This is an important step, as long as it checks the achievability of the bid depending on the availability of the company (equipment, personnel, etc.), constraints imposed by the owner (performance, time, financial factors, risk factors, etc.), costs that are evaluated in relation with the offer and the price which the owner is willing to pay. A decision would be taken at the end of this phase: to continue and proceed to the development of the bid process proposal, or to abandon this bid. In this case, it is necessary that the company capitalizes and validates knowledge and skills built during this inspection phase, which justifies the passage through the closure phase although the company abandons its participation in this bid (Fig.1). In the case where we treat the development of the bid proposition, it is found that the bid is eligible according the vision of the company. A decision would be taken at the end of this phase: readjust the proposal, either finalize and transmit them to the owner, or to abandon this process. In the two last cases, we shall proceed to the closure of the bid process, for the reasons related to feedback and knowledge validated capitalization. Indeed, this phase is devoted to monitoring implementation of the proposal chosen by the owner. Although this choice concerns the works of other companies, the involvement of this phase can only engender benefits for the company at least for its future bid proposals.

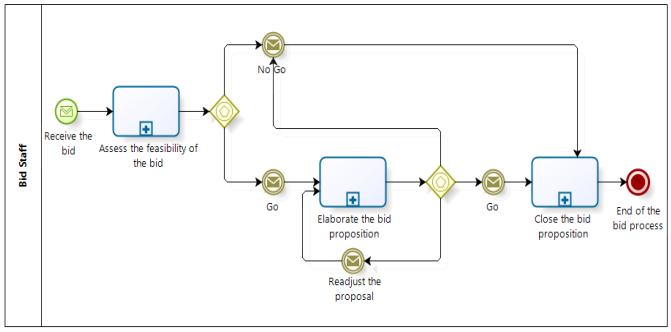


Fig.1 Our Bid Process phases modeled by BPMN.

In the following, we focus on the technical proposal construction. This is a phase that corresponds to the planning of the activity of products designs, which will be subsequently assessed by indicators (technical, business, time, risk, etc.). This is an innovation activity, in so far as it involves human intellectual intervention. Its main objective is to predict the design of products: by optimizing costs of construction, by respecting deadlines of manufacture and by minimizing risks of its future exploitation, of course, while providing the highest quality of services. This requires perfect control of all factors that are critical for a designer, so that he could create realizable solutions, while respecting the constraints and the objectives formalized in the specifications. This activity is submitted to many and various conditions (domain law, manufacturing constraints, etc.). However, the company must mobilize its own human resources to form the bid process team. Practically, it must designate: the bid manager, the cost manager, the risk manager, the expert and the knowledge manager.

At First, experts must identify owner requirements. They should not only provide a list of explicit requirements (expressed by the owner), but they must also be able to reveal the implicit requirements (not expressed by the owner). The requirement specification constitutes a support from which experts could determine the features desired by the owner relative its future product. There should be a list of the main functions, and eventually their sub-functions. Thus experts must weigh each function according to the vision of the owner. In addition, they must specify their criteria for evaluation. After that, each owner function must be translated on product function. For this reason, the experts will be based on the bid memory, specifically on the technical referential, in order to establish a description of product, at a detailed level. It should contain a correspondence memory between requirements and specifications. Therefore, experts choose the most suitable product function, obviously if it exists in the referential; otherwise they will propose transposition alternative on product function. At the same time it would be interesting that the knowledge manager proceeds at acquisition knowledge. Its main objective is to provide validated knowledge to different stakeholders involved in the bid process.

Once the functions of the product are identified, the expert could deduce the type of design to achieve. It is a groovy design of the product, in the case where all the predefined functions were extracted from the technical referential. It is a redesign of an existing product, where the majority of predefined functions are derived from the referential, only a few new functions alternatives have been proposed. In contrast, it is an innovative design for a similar product, where the majority of the proposed functions are alternatives, only a few predefined functions are derived from the referential. Finally, it is a creative design of a new product, if no proposed function was identified from the referential. Once the experts have defined the type design to achieve, we could start building the technical solution. This is the most expensive and important phase of the process elaboration of the bid proposition.

The design strategy defined in the previous step constitutes a starting point for the construction of the technical solution. First, the experts must characterize the product. They must be based on bid memory, specifically on the technical referential, to check if the identified product has already been used by the enterprise, or if he could replace it with a similar product. Otherwise, it is a new product. Then, it should be validated later by the knowledge manager at the cloture of the bid process. Admitting that the main objective of the enterprise is

to measure the technical feasibility and the financial provision of the bid, it would be necessary to quantify the business characteristics of this product. Thus, if we are going to reuse or to readjust an existing product, we should specify the costs and risks related to its use. Otherwise, a decision must be made for a new product, either to buy or make it. In cases where the team opts for purchasing the product, it must be based on the market to predict simultaneously the price of the product and the risks associated with its use. Whereas in the case where it chooses to manufacture the product, it must provide both of the manufacturing cost of the product and the risks associated with its operations. Ultimately the characterization of the product leads to a decision: either by designing the product, or buying it. In the case where the team opts for the design, it should first schedule the activity of designing products and including afterwards the financial proposal. Whereas, when the team opts for purchasing, it must include directly financial proposal.

We show in the following that the ERP (Enterprise Resources Planning) [11] is a support for the design of manufactured products.

III. Required Characteristics and Problems Implantation of the Bid Process Information System

A. Required Characteristics of the Bid Process I.S

The I.S which supports a bid process must be:

Integrated: the company regularly participates in bids, during which, it reutilize its technical skills and business, to subsequently take optimal business decisions. It is in this way that its I.S must be integrated, i.e. capable of creating a comprehensive synergy including (hardware, software, features, and users), so that they cooperate within a single homogeneous system. Integration ensures consistency and harmony on different levels: Business level, Informational level, Decision-making level, and Technical level.

Flexible: the company should survive in an unpredictable business environment: each bid process is realized in a specific context and consider specific solutions. That's why, I.S needs: on the one hand, to be able to overcome the market changes, from one offer to another (adaptability, scalability and extensibility) and on the other hand, to be able to react, on the right time, with the business agility (competitiveness and responsiveness). The company is considered flexible, when it manages to operate, adapt and easily extend its resources at different levels (business, informational, decision-making and technical) to fill and quickly meet the offers' terms during the different occasions of business that it accomplishes.

Interoperable: the company should not address offers and should not conduct business autonomously, and that's why I.S must be interoperable both inside, and outside its functional scope. Obviously, the interoperability is the fact that several systems can operate together while preserving their heterogeneity and autonomy. Thus, internal interoperability is a prerequisite to build an integrated I.S. On the other hand, given the competitive environment in which the business is involved, such a system must foster intercompany cooperation on the spot and in a dynamic way, whenever it is necessary to organize a partnership relation of a bid project, especially, during the construction of the techno-economic proposal: so it requires planning the dynamic interoperability at the external level. Practically, the enterprise is interoperable, unless its I.S is integrated internally and it manages to plan, to synchronize and to exchange: its trades, its data, its resources and its processes, easily with partners from the outer world, and this should happen despite their semantic differences.

B. Implantation of the Bid Process I.S: "three fit" problems

We rely on the urbanization approach to establish our B.P.I.S [9]. This approach is described according to four levels: business view, functional view, application view, and physical view. Nevertheless, this approach is facing the following problems (Fig. 1):

"Vertical fit" problem: the business and functional views are abstract. But, the application and physical views represent definite implementations. It would be difficult to gather the data that allow scrutinizing the company's operations. In fact, this results a fragmentary data in work system and reduces the efficiency of the company. These circumstances prevent from having a system which gives a complete and integrated image on company's inner environment. Integrity, scalability, consistency and transposition are *"vertical fit*" issues that extends from a business infrastructure (logic) to a technical infrastructure (physical) in the company's I.S.

"Horizontal fit" problem: the "*horizontal fit*" translates not only the applications' problems of identification (induced by the "*vertical fit*" problems) that cover the entire infrastructure of the company's business, but the intra-applicative communication problems (internal interoperability) to ensure the interactions between applications of the same technical infrastructure in the company (the same homogeneous system). Such failures make the global system disintegrated and little evolutionary.

"Transversal fit" problem: the "*transversal fit*" translates the problems of inter-applicative communications (external interoperability carried out dynamically).

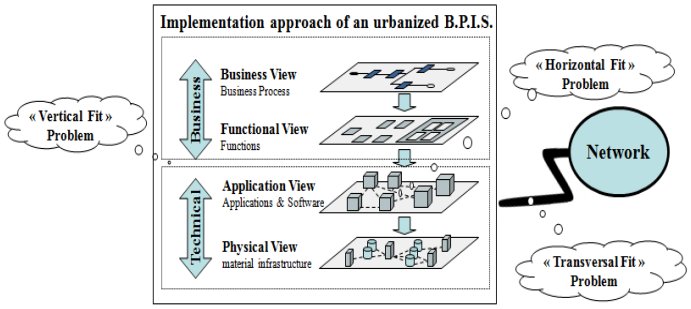


Figure 1. Urban I.S reference model: problems of "three fit" [9].

The "*three fit*" problems prevent the transversal exploitation of bid process, and as a consequence, it is difficult to have a unified bid vision. Solving such problems needs to meet the requirements associated with integrity, flexibility and interoperability on internal as well on external of I.S.

IV. Our Proposition: The Urbanized B.P.I.S

Our objective is to set up I.S which is *integrated*, *flexible* and *interoperable* and which allows a better exploitation of the bid process. We are interested, not only in implementing the right tools to achieve bid in one homogeneous system (integrated), but also, in solving the problems related to interactions intra, and even inter-applicative (interoperability). Our aim is to be able to exploit this system in different bids (flexibility). So far, we dealt with "*three fit*" problems.

We suggest treating these requirements by relying on four essential dimensions: (i) the *operational dimension* that serves to specify the bid exploitation process by undertaking a specific project; (ii) the *organizational dimension* which allow to organize the set of skills and knowledge, that the company acquired during the previous bid in which it participated: the objective is a possible reutilization of this patrimony in future bid projects; (iii) the *decision-making* dimension which aims at optimizing and making the right decisions that concerns the market offers and that takes place during the company's eventual participation in bid processes; and (iv) the *cooperative dimension* which aims at ensuring communication intra-enterprise (internal interoperability) and at planning the inter-enterprise communication on demand , in order to realize a common goal (dynamic interoperability). For example, while creating the offer's technical proposal, one needs to organize inter-enterprise collaborations.

We treat respectively: *flexibility* through the *operational dimension*, *integrity* through the *organizational* as well as *decision-making* dimensions; and *interoperability* through the *cooperative dimension*. In fact, a flexible bid exploitation (*operational dimension*) requires an eventual integration at the level of I.S.; this fact targets reuse the best skills (*organizational dimension*) and adapt the best decisions (*decision-making dimension*), and learn from past bid experiences. All of this requires an internal interoperability within the participating company in the bid process to seek a homogeneous and coherent form of its I.S (*internal cooperative dimension*); and dynamic interoperability, at the level of a virtual company, built to realize in common a bid project, the fact that allows a better coordination and collaboration between various involved stakeholders (*dynamic collaborative dimension*). However, it is the *operational dimension*, which is deemed to be the main focus of our system, which acts and remains sensitive to variations in all other dimensions: the *organizational* and *decision-making* ones. These dimensions are covered by the *cooperative dimension* itself. Indeed, the *operational dimension* takes the *organizational dimension* as a basis in a specific bid. This premise is justified by the fact that the company reuses its own capital of objects for different actions, depending on its needs. This pushes the company create more and more products whose life cycle is shorter than those made in the past. This hypothesis is based on the re-design of existing products or creating similar design products, rather than, on producing new ones. In another way, the *organizational dimension* is based on the *operational one*, both while constructing a bid proposal, or after finishing it. In fact, the former (bid proposal in progress) comes to readjust the proposal, while the latter (bid proposal is finished), comes to update the company's capital through integrating the set of knowledge and skills built during this project. It is worthy to note that this assumption is beneficial for the company's maturity, even when the company abandons its participation in the offer (the company closes its participation before the completion of the bid). In all cases, even if the owner chooses the proposals of other companies (the bid proposal is unsuccessful), a possible updating can take place and influence positively and beneficially, at least for future bids. If we follow the same logic, the *decision-making dimension* relies on the *operational dimension* and vice versa, which consolidate the company's participation in the bid. It becomes evident that the *decision-making dimension* is inexorably related to the *organizational dimension*. All these dimensions are based on the *cooperative dimension*: on the one hand, the realization of the bid can be cooperative (*operational dimension* relies on the *cooperative dimension* for collaborative planning and for the creation of a product); and on the other hand, decisions can be cooperative (the *decision-making dimension* relies on the *cooperative dimension* for a collaborative decision during a definite bid).

We suggest exploiting these different dimensions, while relying on six main approaches, and by describing our urbanization approach of the I.S. Indeed:

The integration of the Lean Manufacturing [12] approach allows designing a product perfectly adapted to the needs of its client i.e. the product can be cheaper in costs but not less efficient than its expected services. Lean adds value to the technical proposal which materializes a bid (Lean participates in solving the problem of "*vertical fit*").

The BPM (Business Process Management) incorporation [13] allows to model business and skill processes, particularly, the bid process. BPM facilitates the alignment of an integrated I.S with strategic directions, regardless technological constraints (BPM participates in the resolution of the problem of "*vertical fit*").

The KM (Knowledge Management) involvement [8] allows formalizing and modelling bid knowledge, whether explicit or implicit, in order to make them operational by during different bid projects that it realizes. KM facilitates establishing a language and covers the organizational dimension within the company (KM participates in solving the problem of "*horizontal fit*"). Furthermore, KM permits improving and responding to individual, collective and organizational learning acquired during a bid process. This hypothesis suits perfectly a bid context (KM takes part also in resolving the problem of "*transversal fit*").

The BI (Business Intelligence) integration [14] allows relying on methods, in order to provide decision-making assistance to those involved in a bid process. Therefore, this approach facilitates implementing explicitly the decision-making dimension within the company (BI participates in the problem resolution of "*horizontal fit*").

The SOA (Service-Oriented Architecture) [4] helps developing an easily flexible, extensible and adaptable I.S

which can be materialized by a set of reusable application components. These application blocks communicate the practical implementation of "services" (clearly defined function in a way that makes it independent of the technical platform). The SOA facilitates the communications' standardization, intra-applicative (SOA participates in the resolution of the problem of "*horizontal fit*"), as well as inter-applicative (SOA participates in solving the problem of "*transversal fit*"). SOA offers an innovative solution to manage the interface between the business needs and its technical implementation (SOA participates in the resolution of "*vertical fit*" problem).

However, the company that takes part in a bid process needs to exploit its "services" at a distance, to promote collaborative work with its partners, such as the work constructed during the technical solution. The integration of Cloud Computing [10], allows the enterprise data and applications to be accessible and usable via the internet (Cloud Computing participates in the resolution of "*transversal fit*" problem).

We can deduce that: Lean Manufacturing, BPM and SOA allow us to overcome "*vertical fit*" problems and thus cover all the dimensions defined previously. KM, BI and SOA allow us to overcome "*horizontal fit*" problems as follows: KM permit to cover the *organizational dimension*, BI can cover the *decision-making dimension*, and SOA enable to cover the *cooperative dimension* at the level of the enterprise. In addition, SOA and Cloud Computing permit to overcome "*transversal fit*" problems and hence cover the *cooperative dimension* to the level of a virtual company. Also, KM participates in solving the "*transversal fit*" problems and consequently cover the *organizational* and *decision-making dimensions*. To this end, it is the *operational dimension* which can be supported by these six approaches. In other words, the satisfaction of this assumption enables us to have a flexible, integrated, and interoperable, I.S something that assures us a better exploitation of the bid process.

We suggest filling these dimensions while relying on the following hypothesis: "ERP (Enterprise Resources Planning) [11] allows us to build the techno-economic proposal of an offer (cover the *operational dimension*), as we rely on the Organizational Memory [6] (to cover the *organizational dimension*). The set of solutions that make this proposal realistic, are going to be evaluated by a Data Warehouse [7] (to cover the *decision-making dimension*)".

V. Operational Dimension of our B.P.I.S.

The ERP [11] aims to pilot the enterprise processes. In this framework, ERP producers suggest integrated packages based on the common skills between enterprises such as: financial, purchase and sales management, production management, technical data management (items, nomenclature, resources, manufacturing process), logistics management, etc. More precisely, these software producers build integrated application modules implementing the market's best practices for each function. Thus, the ERP is the most suitable for the complex processes exploitation within the enterprise, particularly; it helps to exploit the bid process. In fact, the bid process relies on the previously enumerated functions for its exploitation, notably; within the context of making the bid's techno-economic proposal (the ERP covers the bid's *operational dimension*). Nevertheless, there are primordial functions for the bid process evolution, but these functions are not treated by the ERP such as: expertise management (the ERP does not cover the *organizational dimension*), risk management and decision-making management (the ERP does not cover the *decision-making dimension*). This justifies our choice, in the previous paragraph which states to support the ERP by other tools in a way that enables treating a bid process. In fact, the ERP targets to meet the bid process *operational dimension.*

A. ERP: the heart of an industrial I.S operational dimension

With the intention to manufacture valuable products to its customers, the manufacturing enterprise must plan the use of its resources (raw material, labor and machinery) while respecting the set of practices, rules, tools and methods. Such requirements represent the enterprise industrial culture. In other words, the company's skills and expertise enable to ensure a good exploitation of its resources and more efficient production as expected by its clients. For this reason, the manufacturing factory is based on three tools' categories (piloting, support, and production). These tools meet the necessities of both customers as well as partners. Subsequently, we show that the ERP is at the heart of the industrial I.S., acting and remaining sensitive to variations of all the other tools making this system (the ERP is a piloting tool which supplies and provides the interface between support and production tools). We suggest describing the *operational dimension* which meets the requirements of an industrial I.S. by the following tools:

CRM (Customer Relationship Management): is a system that manages customers' relationships, hence, it supports all the processes organizing the company's interactions with its customers (sale, marketing, and after sale services). In fact, while realizing a specific bid process project, the ERP disposes a module that allows managing the relationships with customers. However, it can rely on a CRM to check if the owner appears in his clients' list. This allows the company to view the history of its relationships with its customers in order to highlight their preferences. Such a strategy grants to increase the clients' loyalty and satisfaction, hence, it increases the enterprise chances to win the bid. In all cases, the CRM is a support tool that allows the company to have a set of coherent, reliable and accessible data for the client and as a result it delineates a better picture of their interactions.

PLM (Product Life-cycle Management): is a system that manages the technical data related to products' designing and manufacturing. It is noteworthy that the ERP is equipped by a module that aims to manage the technical data, and even to plan the product construction. As a result, in the case of a specific bid process, it can cover the realization of the technical proposal which makes the offer. However, the fact that the ERP relies on PLM enables to reuse and adapt the processes and the existing equipment, and to visualize how they have been exploited in past projects. This strategy allows the company to operate its skills and technical abilities within a bid process project. This fact enables to optimize time and design costs of the technical proposal and to increases the chances of winning the bid. PLM is a support tool that can even maintain the standardization and coherence between the procedures used by different participating companies in the project: it facilitates the organization of data on the product.

SCM (Supply Chain Management): is the system responsible for the management of the enterprise logistic chain and its relationship with its partners (suppliers, subcontractors etc.). In fact, as soon as the client makes his order, the company and its employees start to find supply sources and seek to deliver the product or service on time, at lower cost according to the choice of the applicant and at a reasonable price. Thus, the SCM is a support tool that allows managing the physical flows of the company. It allows automating the movement of information flows beyond the borders of the company. The SCM helps the company to create relationships with partners and spread its work in other companies. Note that the ERP has a module that allows planning the supply chain, in the context of a specific bid process. Once the bid is won, the SCM relies on ERP to follow and carry out, onsite, the plans done by the company and its partners.

MES (Manufacturing Execution System): is a system designed to obtain instantaneous supervision of a production workshop. It manages manufacturing orders coming from the ERP and sends back the state of outstanding amounts and a report on the events that occurred in the workshop. It is a tool supervising production and aiming to control the operation of the set of resources (materials, equipment, staff, etc.). It offers the company's users analysis indications on the equipment performance, as well as some pieces of information that describe the production status (work in progress, finished products, machines, etc.). Note that the MES allows actors, who are involved in the production, to work with a unique and homogeneous system even when they are from different departments of the same enterprise.

The SCADA (Supervisory Control and Data Acquisition): is also a control system dedicated to production. Specifically, it intends to pilot and automate the production processes. This system favors the acquisition of monitoring data related to the execution of these processes environment. From an MES viewpoint, the SCADA system constitutes a reinforcement to manage production.

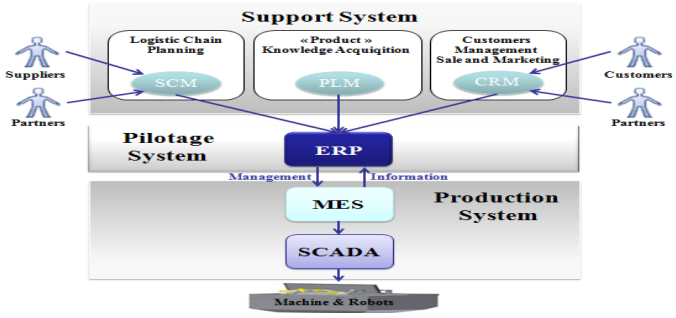


Figure 2. Operational Dimension of an Industrial Information System.

We can deduce that SCM, CRM, and PLM are support tools, while MES and SCADA are production tools. Fig.2 reveals that ERP is at the heart of the *operational dimension* of an industrial I.S. However, the enterprise that produces goods and services exploits these different business applications, without really integrating them, to follow the value chain. This causes consistency problems at the very heart of its I.S. it is also a constraint to manage flows, tasks, order and batch, etc and their corresponding information in real time. Actually, the main failure concerns intra-applicative information interactions and exchanges, especially between the piloting tool (ERP) and the producing ones (MES and SCADA). This limitation causes communication problems between business and manufacturing levels in the company which uses different software's (problems of terminology sharing between these different applications, which cause internal interoperability problems).

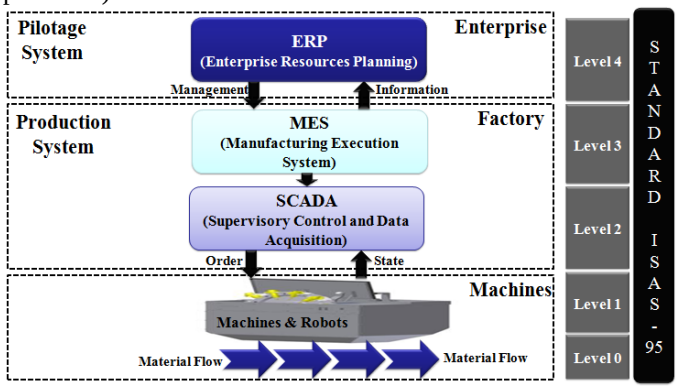


Figure 3. ISA S-95 Integration standard ensuring coherence at the level of the operational dimension of an industrial S.I..

Note that, the norm ISA S-95 [2] suggests mitigating this limitation by defining terminologies and information models in order to establish a coherent communication between these two levels. Thus, integrating these norms can establish coherence between: ERP which ensures a good management of resources; the MES which aims to achieve an effective supervision of the manufacturing system; the SCADA system which targets to optimize the supervision of processes and the acquisition of data. It is noteworthy that the MES is a bridge between the piloting system process and the organizational management level since it combines ERP and SCADA systems. In fact, the ISA S-95 standard intervenes in different levels (Cf. Fig. 3): (level 0) represents the current production processes; (level 1) defines the activities related to physical processes; (level 2) ensures the physical processes supervision and control; (level 3) defines the activity flow to manufacture finished products; and (level 4) defines the activities related to the company's business. These activities are necessary to manage the manufacturing organization.

B. Our Positioning Regarding ERP evolution

During the past two decades, the ERP(s) strongly evolved depending on the progress of NICT (New Information and Communication Technologies). This evolution is described in terms of three successive generations [3].

The 1st generation ERP intends to make synergies between the company's resources by integrating operational processes and their information flows. The presumed ERP model is considered as a management system adapted only inside the enterprise. Concretely, it aims to control costs and to develop specific transactions in order to establish exchange coordination between partners.

The 2nd generation ERP becomes interested in platforms' design and reengineering integration, internal versus external, business processes. These platforms are aligned to business strategies which are reputed more and more collaborative and dictated by the activities as well as the market. As far as its interactions with the environment, market and partners are concerned, the presumed ERP model is more dynamic than the previous one. Thus, the ERP model developed out of the previous transactional approach, towards a classic relational approach that is based on closed exchanges between partners of the same chain of value (dubbed "ERP pipeline").

The 3rd generation ERP is interested in developing a management system that integrates in skill processes through the alignment of ERP business processes. Thus, the ERP model evolves from a "closed," relational approach towards an open one including a multitude of business networks. This model is based on the most innovative network architectures mobilization, supporting interoperability and introducing: web services, e-market and online business models.

This work contributes to the 3rd generation. In fact, the bid process requires the exploitation of inter-enterprise cooperation on demand. This consists of interconnecting a heterogeneous set of processes that aim to achieve a common goal, for instance, the case of preparing the technical proposal of the offer. This interconnection of processes must be carried out according to a definite approach while respecting the constraints mentioned in the specification. Therefore, the 1st generation ERP which is limited to the management process that covers only the internal borders of the company does not meet this objective. The bid process is also known as a transverse, dynamic, open and not closed process. As a consequence, the 2nd generation ERP does not respond to this objective as well. This is justified by the fact that the new business model is strategically and operationally directed towards researching agility in the modes and processes of management. This model lies on a hard competition in an open business network.

C. ERP at the Heart of Coupling the Technological Capacities and Management

The success of implementing an ERP project dedicated to a particular enterprise must pass by design and platform coupling of reengineering and integration, internal versus external, trades and business processes. We are facing an exigency related to a specific governing strategy of each enterprise on each trade. This strategy intends to establish a simultaneous correspondence and alignment between two distinct objectives: the first refers to the ERP agility (the ability to adapt to different management modes and to harmonize its exchanges with the appropriate partners). However, the second objective is linked to management processes, particularly, those exploited by a bid process (the enterprise's ability to innovate in an incremental and dynamic way opens the ground for a better management of costs, deadlines, risks, etc.).

Our goal is to set up a 3rd generation ERP ad-hoc and its strategy as a solution. Therefore, the technical capacities and management coupling impose themselves. However, adaptability to the context or "conduit change" is a major factor of success for an ERP project throughout a functional integration. It is, on the one hand, a "governing affair" conduit which assures innovation in the management of business and trade processes. On the other hand, it is a "governing technology" conduit which ensures the dynamics of these processes. Besides, studies formalizing the combination between the two previously mentioned objectives are rare or non-existent. It is not easy to deploy a solution that aims to respond to such governance, combining an evolutionary and incremental approach to implement a technology that seeks the company's I.S. agility while allowing a better management of its business and trade processes. As a result, we rely on the NICT to answer this problematic. ERP has some limitations on management and technical levels. As a consequence, since the ERP relies on NICT, we are able to meet the limits detected on those two levels. This solution enables us a better management of the bid process. It is true that the ERP aims to manage the bid process and its information flow. However, the fact that the ERP relies on a BPM approach, this allows to model the bid process and to subsequently align the ERP on strategic directions, without undergoing many technological constraints. Such a strategy helps the bid process standardization (this allows to have different ERP(s) with different layers of trade standards while realizing a bid process). Hence, the cooperation between enterprises and more accurately between ERP(s) (let us remember that an ERP covers a bid process *operational dimension*) is easier to realize, and in this way we get the impression that we are working on the same ERP (this solution is highly recommended during the collaborative construction of the offer's technical solution). Consequently, the BPM [13] will help us to construct an ERP in a modular way by assembling trade components weakly coupled. It remains only to solve the communication between different technical infrastructures which materialize the different ERP(s). We previously showed that the SOA [4] allows solving such problems.

The ERP enables to align processes best practices. However, it is unable to follow these good practices. Thus, the ERP does not adapt with the continuous improvement of the enterprise in its affairs. The fact that the ERP relies on Lean Manufacturing [12] strategy enables to overcome this limitation and helps the enterprise to establish a culture and a permanent maturity that revolves around good practices during the different contributions in different bid processes.

The ERP does not permit reusing the enterprise's expertise acquired due to previous bid processes. As a result, the ERP must rely on an external system that enables to collect and to restitute the data and skills, in order to make them available, operational, and exploitable by an ERP during a specific bid process, in real-time. Therefore, the fact that the company relies on a KM approach and more accurately on an OM (Organizational Memory) [6], this enables the ERP to effectively exploit the enterprise's internal language in different bid processes. Note that, we suggested in [15] an OM sustained by an ontological framework in order to operate on certain business processes, we can accommodate this to realize a context for a bid process.

The ERP is designed to collect the event traces, but it is not created to help the decision-making process. Consequently, the ERP inability to treat uncertainties and unexpected events limit its use to support the decision-making process in an environment of dynamic production. Therefore, the fact that the ERP is based on BI approach [14], and more precisely on a DW (Data Warehouse) [7], enables the enterprise's employees to get a rapid and synthetic access to strategic information. Furthermore, these employees can easily adapt their decisions taken in a past project to a specific bid process project.

We can deduce that a coupling between an ERP and management capacities (BPM, Lean Manufacturing, KM et BI) are highly recommended to meet the "governing affairs" level of a enterprise in a bid process project. Following the same logic, we demonstrate in what follows that a coupling between ERP architecture and technological capacities (SOA and Cloud Computing) is absolutely necessary to meet the "technological governing" level of the company in a bid process project. The structures provided by an SOA allow implementing ERP ad-hoc architecture in the enterprise hence helping it to surmount environmental turbulences, to respond to agility affairs quickly, and to improve the management of its markets and its bid processes [3]. Moreover, ERP solution based on SOA architecture helps to develop, at the very heart of the enterprise, new skills thanks to reutilization [4]. This strategy enables to enhance the company's expertise and to establish, as a consequence, a permanent culture of change. This is revealed to be necessary for the company's eventual adjustment in different contexts. These characteristics strongly favor the management by processes (this is strongly recommended to be able to manage and exploit the bid process). In fact, the architectures suggested by SOA are flexible, adaptable reusable and extensible to a great extent. These architectures enable to easily integrate the new affairs costs realized by the enterprise in a specific bid process and to constantly develop trade and business processes modeling. Furthermore, the ERP deployed with SOA architecture ensures a low intra-applicative coupling (low coupling between ERP and the other applications used in the enterprise) as well as a low inter-applicative (low coupling between different ERP of different enterprises to realize a bid process).

In certain cases, the ERP needs to exploit distant functions which are hosted externally. This requirement is realized if the enterprise includes Cloud Computing [10] approach during the implementation of its applications. Thus, the ERP must interact with applications of SaaS (Software as a Service) [5] mode each time it needs to exploit new functions which it does not cover. This solution is strongly recommended during the collaborative construction of a bid process proposal.

Finally, we validate the hypothesis that we departed from: *"the fact that the operational dimension ERP is based on six approaches (BPM, Lean Manufacturing, OM, BI, SOA and Cloud Computing), this helps us to meet other dimensions (organizational, decision-making and cooperative). Such solution allows us to have a flexible, integrated, and interoperable I.S., which helps us to ensure a better exploitation of the bid process."*

VI. Conclusion and perspectives

We presented our methodology of implementing a B.P.I.S. We showed that such a system must be integrated, flexible and interoperable. However, during the implementation of this system, "*three fit*" problems (vertical, horizontal and transversal fits) fail the inclusion of such requirements. To overcome these deficiencies, we proposed to address the I.S design following four dimensions: the *operational dimension* (tackles flexibility), the *organizational* and *decision-making dimension* (tackles integrity) and the *cooperative dimension* (tackles interoperability).

We showed that the ERP is the most suitable tool for the exploitation of a bid process, thus, it is a constructing support techno-economic solution of the offer (the ERP targets *operational dimension* of a bid). We demonstrated that the ERP should be used at the heart of coupling the technological capacities and management.

Our work opens the horizon to exploit this suggestion within a concrete bid project.

References

[1] Alquier, A.M, Tignol, M.H. : Management de risque et Intelligence Economique, l'approche PRIMA, Economica, Paris (2007).

[2] ANSI/ISA. Enterprise Control System Integration - Part I: Models and Terminology for Enterprise-Control System Integration. Instrument Society of America (2007).

[3] Baile S., Eddir H. : La conduite du changement au cœur du couplage des capacités de la SOA et du BPM- Contribution au schéma conceptuel de l'alignement d'un ERP à la stratégie d'affaires d'un PME", 17ème congrés de l'AIM, bordeaux, (2012).

[4] Bean, J.: SOA and Web Services Interface Design, Principles, Techniques and Standards, Elsevier, Burlington, USA (2010).

[5] Cancian, M.H., Rabelo, R.J: Supporting Processes for Collaborative SaaS, IFIP Advances in Information and Communication Technology, Vol. 408, 183–190 (2013).

[6] Connolly, J.: Corporate DNA: Using Organizational Memory to Improve Poor Decision Making, Journal of Management History, Vol. 16 Iss: 1, 137–138 (2010).

[7] Dahiya, N., Bhatnagar, V: Effective data warehouse for information delivery: International Journal of Networking and Virtual Organisations, Vol. 12, N° 3/2013, P 217–237 (2013).

[8] Dalkir, K.: Knowledge Management in Theory and Practice (2nd Edition). Cambridge (Massachusetts): MIT Press, London (2011).

[9] Fournier-Morel, X., Grojean, P., Plouin, G., Rognon, C. : SOA le Guide de l'Architecture du SI. Dunod, Paris (2008).

[10] Gerald, K.: Cloud Computing Architecture, Corporate Research and Technologies, Munich, Germany (2010).

[11] Gronau, N.: Enterprise Resource Planning. oldenbourg.verlag, 2nd Edition (2010).

[12] Skaf K.M.: Application of lean techniques for the service industry. Technical report. A case study M.S., Southern Ilinois University at Cabondale (2007).

[13] Weske, M.: Business Process Management: Concepts, Languages, Architectures, Leipzig, Germany: Springer-Verlag, Berlin (2007).

[14] Yu, E., Horkoff, J.: Business Modeling for Business Intelligence on Data Management In Perspectives on Business Intelligence, Vol. 5, N° 1/2013, 19–32 (2013).

[15] Zahaf, S., Gargouri, F. : Mémoire organisationnelle appuyée par un cadre ontologique pour l'exploitation des processus d'affaires: 31ème Congrès en INFormatique des ORganisations et Systèmes d'Information et de Décision INFOSID, Paris (2013).

A secure E-Court System for India

Himanshu Agarwal
IEEE MEMBER
Dept. of Information Technology
IIIT-Allahabad
India-211012
agarwal1488@gmail.com

Harshal Singh
Dept. of Information Technology
IIIT-Allahabad
India-211012
harshal90singh.hs@gmail.com

G.N.Pandey
vice chancellor
Arunachal University of studies
Arunachal Pradesh, India-211012
Adjunct Professor, IIIT-Allahabad
pandey63@gmail.com

Abstract: The main motivation behind writing this paper is the pendency prevailing in Indian jurisdiction system which is about 35 million as per the survey. The technology of E-court for India is proposed for the first time in this paper. This proposed model has a greater security due to unique identity with their finger prints. Considering the proposed model for e-court, the criminals will be provided hearing opportunity from the jail itself through Audio/Video facilities. This proposed model is also eliminating the possibility of their running away while under transit from jail to court. The proposed model will be extremely helpful in early disposal of the court case.

Keywords: Biometric device, E-court, Online judicial system, Finger print Recognition,

I. INTRODUCTION

E-court is the process of making the complete existing Indian Courts online. In the detailed description of E-court, it could be understood as the process of making the court proceedings electronic with the help of audio/video conferencing, it also allows the linking of hospitals, forensic labs, police stations and prisons. Evidences, affidavits, FIR reports and unique identity could be submitted online as and when required. User could know about the lawyer and could choose one and pay the fees online. Finally the status and the result of the cases could be shown online.

The project aims to build a national grid of key judicial information available 24 X 7 in a reliable and secure manner. E-Court, an integrated MISSION PROJECT, has a clear objective - to re-engineer processes and enhance judicial productivity both qualitatively and quantitatively to make the justice delivery system affordable, accessible, cost effective, transparent and accountable.

The scope of E-court is to develop, deliver, install and implement automated decision making and decision support system in courts all over the country.

The main objectives of this system are:

- To assist judicial administration in reducing the pendency of cases.
- To provide transparency of information to the users.
- To provide access to legal and judicial databases to the judges.

In this proposed model prisoner could attend hearing from his own place/jail with the help of audio/video conferencing. It is more secure as compared to the

present system. It also helps in achieving timely justice.

This proposed model also helps in the work of online depository like preparation of online affidavits, online deposition of stamp fees, online registry of land, houses and vehicles and related issues.

The evidences could be uploaded online; the whole proceedings could proceed electronically. Finally the proposed model also shows the results online and aids in a faster delivery of justice.

This proposed model is more secure, fast, efficient and also increases the transparency of the court proceedings.

II. SYSTEM OVERVIEW

The proposed system is approachable from two sides 1st from the court side (judge and advocate/administrator) and 2nd from the user side.

A. Registration

First of all the interested user will open the login form for registration and he/she will fill all the details therein and after that he/she will be registered. We are proposing a method in which user will have to upload the FIR Report (wherever required) and Unique Identity while registration for additional security, only after that user will get registered. This will allow putting a check on any random user to get registered and only genuine users will be registered to the E-court.

B. Verification

The user will be registered after being verified by the administrator. The administrator will register the genuine users based on their registration form.

After verification server will generate the random username and password, and these will be send to the registered phone number and email of the user.

C. Login

The user will fetch the username and password from the registered phone number or email and with the help of this username and password the user could login to the e-court. Once the user has successfully logged in, he/she will be redirected to the home page, there he/she could:

- Know about the e-court: which includes various attributes of the E-court.
- He/she could choose the lawyer from the lawyer profile page, corresponding to the types of cases.
- He/she could file a case and could view a case.
- He/she could edit his profile.

Input / Output scenario of E-Court

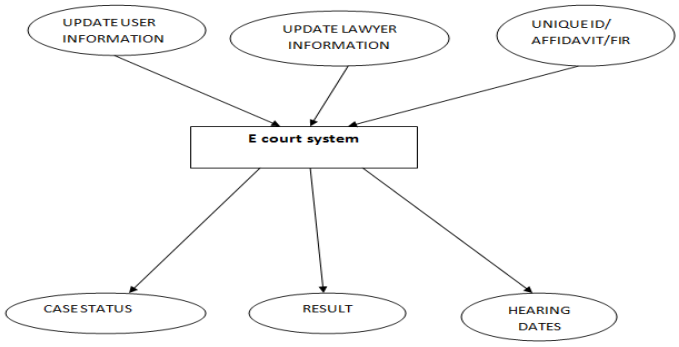


Fig.1. Input / Output scenario of E-Court

D. Case proceedings

In this paper we are proposing a method in which the user will have the complete authority to choose the lawyer online and could pay the fees of the lawyer online. The proceedings will start by filing the affidavit by the advocate of the first user online

and this will be open for viewing by all citizens of India. The advocate of the person against whom the case is being filed will file the counter affidavit in response to the first affidavit, and this will also be done online. This shows the involvement of both the parties in the same case. In response to the counter affidavit from the second user the first user will file the online rejoinder. The administrator only will have the authority to approve these affidavits. All these documents will only be in read only mode so that no one could modify it. If in case these documents need some modifications then this could be done only with approval of the administrator. The administrator will have the complete authority either to accept the request for modification or could reject the request for modification.

For the sensitive cases, the proceedings will held only in the presence of related people and no one apart from the related people will be allowed to witness the case proceedings.

The case proceedings will be done through the audio/video conferencing aids so as to avoid the wastage of time and to ensure the timely delivery of justice because the users will be able to attend the proceedings from all the related places such as police stations, hospitals, universities law firms, jails etc. Whole data will be stored online and could be retrieved later at any time from any location.
This also proves as an example to avoid wastage of time. Another proposal is linking of hospitals, forensic labs, and police stations and providing higher level of security. Finally the status and the result of the cases will be displayed online by the administrator.

Detailed Architecture of Login Process

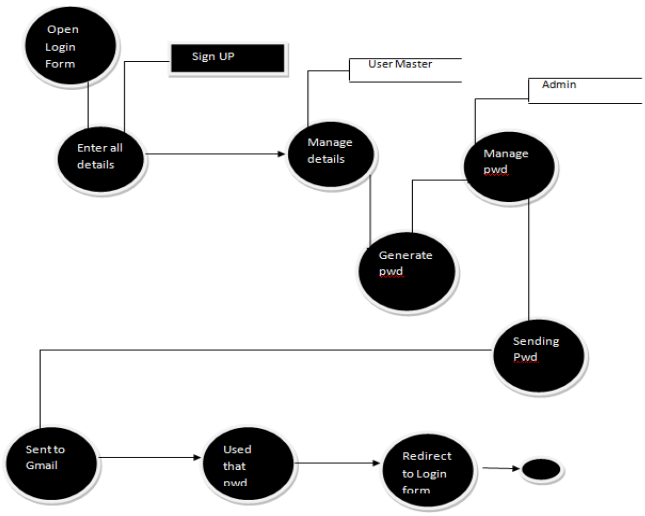


Fig.2. Detailed Architecture of login process

III. MODULES

There are four different modules in the proposed model. These modules work in synchronization with each other for the proper functioning of the project. These modules are as follows:-

A. Administrator Module

The administrator of the E-Court is responsible for sending password corresponding to username to the user's registered mail id and phone number so that user could login securely. Administrator will also update and maintain the lawyer profile. Random case id generation for every case a user has filed will also be done by the administrator. Administrator will also maintain and update the various databases. Administrator will also update the information regarding the hearing dates. User and lawyer authentication will also be done by the

administrator. Finally the case status and the result of the cases will also be shown by the administrator.

B. User Module

User could login into the E-Court with the help of the username and password being sent to the registered phone number and email id. After login the user will search the lawyer from the lawyer profile page, where he will be able to see various lawyer details such as lawyer unique id, lawyer name, total number of cases fought by the lawyer, total number of cases won by the lawyer, lawyer fees, lawyer contact details, and lawyer educational background. Based upon the convenience the user will select any lawyer by sending the request to corresponding lawyer which is subjected to approval or rejection by the lawyer itself. Apart from this user will have an option to update his personal details.

Process of selecting Lawyer

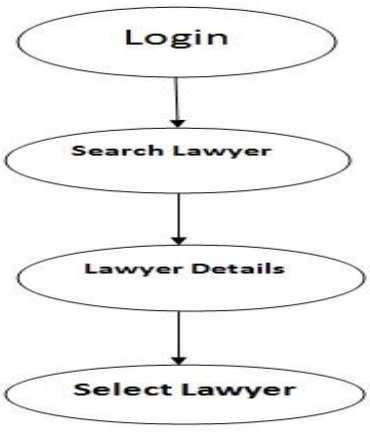


Fig.3. Process of selecting Lawyer

Lawyer's Profile Page

Lawyer UNIQUE ID	Lawyer NAME	IMAGE	Lawyer PROFILE	Fees	Contact	EDUCATIONAL DETAILS	No of Cases fought	No of Cases won
2343HJDSJ86	H.Singh		▢	25 k	534232	▢	76	62
6545JSLSN56	S.Raghavan		▢	40 k	234235	▢	456	370
8975KSMOK56	N.Mishra		▢	74 k	643533	▢	754	700
9876DKLKZ09	P.K.JAIN		▢	24 k	532456	▢	455	408
2435LSKSK54	A.DAS		▢	56 k	645534	▢	75	66
8643LSKKN76	M.H.ROY		▢	62 k	543523	▢	45	33
9863SDJCH95	R.SHETTY		▢	84 k	435674	▢	675	603

Fig.4. Lawyer's Profile Page

C. Lawyer Module

Lawyer is responsible for updating his/her personal profile which includes name, contact details and fees only, other details will be managed by the administrator like lawyer identity, number of cases fought, number of cases won. Lawyer will have the authority to accept any case or reject any case. Lawyer will be able to upload the evidences which includes FIR Reports, affidavits, counter affidavits, rejoinders etc.

D. Notification Module

Notifications will be provided to the user regarding their passwords and hearing dates with the help of a mailing engine. Time to time the status of the cases and final result of the cases will be notified with the help of notification module.

Detailed Architecture of E-Court

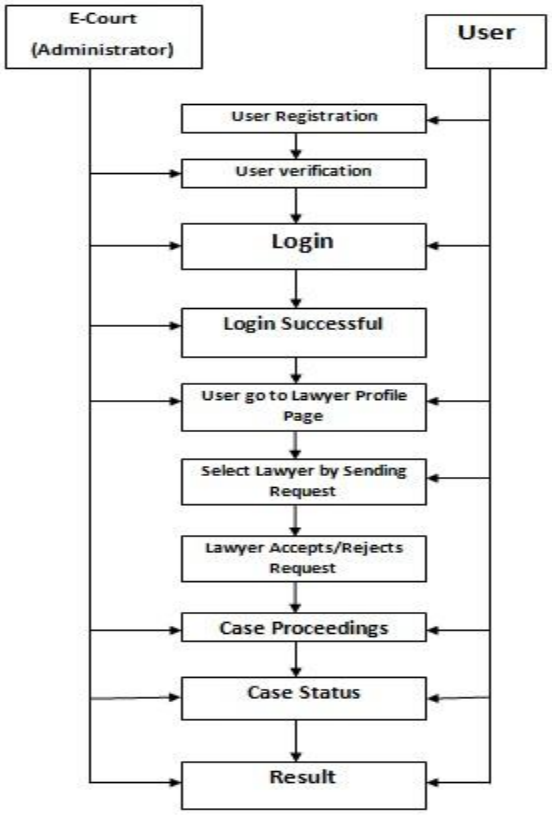


Fig.4. Detailed Architecture of E-Court

IV. FEATURES AND ADVANTAGES

A. Less preparation

Lawyers preparing cases could increase the speed in which documents are trial-ready and decrease paper costs.

B. Cost savings

It is estimated that the reduced time taken to process cases through the courts and the consequent cost savings is up to 30 percent. In addition, the system could be used outside the boundaries of expensive courts.

C. Shorter trials

The cases due to sit for a long time or those that have a number of complex issues to review, could be presented much more succinctly. Any documents could be instantly displayed on screens around the courtroom, without the loss of any time trying to locate the right physical document or file.

D. Clearer evidence

Ready-made functionality, like comparing signatures or displaying the market share price, aids explanation in complex cases. By highlighting relevant evidence in pictorial or diagrammatical displays, counsel could also focus the court's attention on the legal point being made, increasing the impact of their arguments and making it much easier for the judge and jury to understand.

E. Greater revenue

Counsel and case teams are able to process more cases through the judicial system, maximizing their revenue.

V. LIMITATIONS

A. Power Backup

Due to budgetary constraints, funds for providing connectivity from diesel generators sets to ICT (Information and

Communication Technology) Infrastructure may not be provided under the E-Courts project. All such cabling and installation costs have to be borne by the High Courts. Recurring cost for maintenance of diesel generator sets shall be borne by respective High Courts (highest courts in any state only next to Supreme Court in India).

B. Project Monitoring Website

There have been delays by High Courts (highest courts in any state only next to Supreme Court in India) in uploading the required details on the website. Since NIC (National Informatics Center) is dependent on the website for key project data, all High Courts must ensure that data uploaded is complete and accurate.

C. Videoconferencing Facility

The project envisages video conferencing facility at 500 locations. Providing facility only at district courts without covering jails will not be beneficial. However, additional resources are essential for providing full and effective connectivity between jails on one hand and civil courts on the other for timely disposal of criminal cases in India.

D. Inadequate funds for Site Preparation

Cost estimates for site preparation in the original proposal were insufficient, due to increase in number of court complexes. Hence, additional funds are required for site preparation.

VI. CONCLUSION

It is, thus, obvious that when the concept of E-court is implemented, pendency in the Indian judicial system will be slowly minimized and there is every possibility that with exponentially increasing connectivity with proper bandwidth, Indian judicial system will be further strengthened to eliminate any delay in the judicial system.

REFERENCES

[1] http://www.prsindia.org/parliamenttrack/vital-stats/pendency-of-cases-in-indian-courts-2011-1837/

[2] http://theglobaljournals.com/paripex/file.php?val=MTgzNA

[3] Agarwal, Himanshu, and G. N. Pandey."Online voting system for India based on AADHAAR ID." ICT and Knowledge Engineering (ICT&KE), 2013 11th International Conference on. IEEE, 2013

[4] "Supreme Court chides itself, govt for judicial backlog". Times News Network. 12 Jan 2012. Retrieved 11 January 2012.

[5] "Governance in India". worldbank.org. Retrieved 21 January 2012

Analysis and Impact of Bias in Market Research

(A behavior-centered analysis)

Edmund Balogun
The City University of New York, City College

Abstract - *There is no doubt that today almost every product or service available to consumers, retailers and even producers undergo some form of market research and testing. Market research is an essential part of any company or organization, as it enables the firm to understand specific areas to focus on during the development and launch of a new product/service. While market research has been seen to improve the quality of end products and increase product acceptance amongst consumers, there is still an overlapping gap associated with market research. Often categorized as Bias, this overlap tends to alter the results of research and sometimes disrupts the outcome of a product/service. This paper highlights some biases associated with market research, it provides insights to how bias can affect research results, and finally it proffers some best practices that can be adhered-to by industry experts.*

Keywords: *Market research, research bias, Biases*

1 Introduction

Market research is usually carried out in multiple ways and several techniques are utilized. The most common techniques are surveys, focus groups and interviews—these techniques are effective in gathering insights, attitudes, behaviors and perceptions of consumers; however they tend to introduce some form of bias. Bias in research has been defined as "any force, tendency, or procedural error in the collection, analysis, or interpretation of data which provides distortion" (Tortolani)—therefore bias can affect the outcome of a given study. An impending cause of bias in market research is that occasionally researchers use a single technique, mostly surveys in conducting a study. A drawback of using the survey model independently is that it introduces Common Method Variance (CMV). CMV can be described as "the inflation, or in rare instances deflation, in the true correlation among observable variables in a study." (Naresh K. Malhotra, 2007).

Some market researchers often acknowledge the presence of these flaws, and they constantly strive to limit biases and make adjustments as much as they can—but there have been reports that some market researchers tend to induce bias intentionally as a means to uncover findings that may not be revealed by conventional means (Tortolani). Though this method might yield some positive results, it can be unethical and can distort a study completely if uncontrolled. The main focus here is to underline biases often associated with focus groups, surveys, and interviews, as well as provide instances of how they can affect results.

2 Discussion

As mentioned earlier, there are multiple ways in which bias can be introduced in a study. Regardless if it's a survey, a focus group or interview, the tendency for the results to be somewhat biased is relatively high—depending on the conditions in which any of the mentioned techniques' are employed. To some extent it is arduous to categorically list all types of bias in research, however the most common types of biases in market research are method bias, sample bias, self-selection bias and confirmation bias. It is relevant to note that these biases span across sub-categories and sometimes are addressed as such. Certain researchers have recognized shortcomings in their efforts to mitigate bias in market research, as *Roger Dooley*, a Forbes contributor states "when we interpret customer feedback, or data from surveys and focus groups, it's our natural tendency to interpret the data in a way that is consistent with what we believe ourselves. As the data rolls in, we want to blurt out, "I knew it!" or, "I told you so!" Rarely do we look for other ways of viewing the data, particularly explanations that might prove us wrong" (Dooley, 2013).

While misinterpretation of data is an underlining cause of bias in market research, there are other causes like "*the interpreter in your head"*—a notion that people are often incapable of articulating why they do things or how they would behave in the future (Dooley, 2012). Improper wording on questionnaires is also a cause of bias. It is important to state that bias in research is different from errors. There are several types of errors that can occur in a research during data collection and analysis. Examples are sampling error, nonresponse error, and measurement error—the focus here is on types bias, thus it is of little significance to further elaborate on these errors.

2.1 Method Bias

Method bias is very common in market research; it is an empirical cause of measurement error. Method bias or method variance as it is sometimes called "refers to variance that is attributable to the measurement method rather than to the construct of interest" (Philip M. Podsakoff, 2003). The existence of method bias can pose difficulty in grasping the actual phenomenon that researchers are studying because it

changes the actual associations that would be present with the variables (Naresh K. Malhotra, 2007). Sources of method bias includes but are not limited to social desirability, positive and negative affectivity, and consistency motif (Naresh K. Malhotra, 2007). Method bias can occur when respondents provide answers that are socially acceptable instead of providing more accurate and truthful answers—respondents might understand a question in a different way not intended by the researcher. Method bias is hugely present in surveys and it can be damaging to the final results of a study.

2.2 Sample Bias

A sample is a group of research participants selected to represent a larger group or population (Huffman, 2010). Sample bias often arises from not selecting a truly random sample that is representative of a larger population—it is a systematic difference between the groups being studied. For instance, "much research has been done on the increased safety of having air bags in automobiles. Unfortunately, however, the research has been conducted almost exclusively with men. When some car manufacturers apply findings from this research, with no regard for the sample bias, they create air bags sized for men. Tragically, in the event of a crash, these male-sized bags may seriously damage (or even decapitate) small adults (mostly women) and kids" (Huffman, 2010).

2.3 Self-Selection Bias

Self-selection bias occurs when survey respondents are allowed to decide entirely for themselves whether or not they want to participate in a survey (Olsen, 2008). Self-selection bias is almost unavoidable, as researchers cannot necessarily force participants to respond or participate. Some respondents may be less likely to respond for various reasons; others might be uninterested or utterly unwilling to participate. For example, a school program might record high scores on its evaluation and credit the program as excellent but in reality not everyone in the program actually completed the evaluation questionnaires.

2.4 Confirmation Bias

Confirmation bias is especially common in focus groups and interviews, but not excluding surveys as well. As (Dooley, 2013) puts it, confirmation bias is "the tendency that influences all of us to put more faith in information that agrees with what we already believe, and discount opinions and data that disagree with our beliefs." Take for instance, if a researcher believes that price is the most important consideration in respondents' purchasing choice, the researcher might be tempted to structure questionnaires in a misleading way—e.g. "How important was price in your purchasing decision?" (Petri, 2013) in such conditions, respondents' are almost certainly going to reply "very important."

3 Recommendations and Best Practices

Efforts have been made by clinical psychologists to curb biases associated in medical research—one effective method used in drug testing research is administering of *placebos* to research participants. A *placebo* is a fake (inactive) drug given to participants during a drug testing research; it is used to test the effect of the *placebo* on the participant as opposed to the effect of the actual drug administered to a different participant (Huffman, 2010). Very often, a *placebo effect* is recorded during such studies; participants who received the fake drug without any knowledge of it believed the drug relieved them of whatever condition they might have had. This model has been seen to reduce bias in drug testing research. Conversely, in marketing research, there still arises some controversy about the best ways to approach respondents, or the most effective techniques to utilize in a market research.

Completely eradicating bias in market research is a farfetched notion; nonetheless researchers are finding ways to mitigate causes of bias in research. *Roger Dooley*, an advocate for *neuro-marketing*—a model that analysts' use to study the pattern in which the brain responds to sensory and cognitive marketing stimuli; suggests that researchers should develop a more holistic approach when conducting a research. He also recommends that researchers adopt a different outlook and a broader perspective when analyzing certain situations—an example he explains is adopting Warren Buffet's approach to circumventing confirmation bias. An approach in which Buffet first acknowledges that his decisions could be influenced by confirmation bias, and as such he carefully takes into consideration ideas of his critics' which completely contradict his own.

Measuring respondents' actual behavior as opposed to asking them is also an effective way to reduce bias—this can be achieved by adopting the *neuro-marketing* model proposed by *Roger Dooley*.

Furthermore, *Dooley* recommends that researchers look for alternative means to interpret information, while paying attention to nonconforming opinions within the team or company when analyzing customer feedback. Adopting new ways of working differently like employing creative thinking, engaging participants, talking to specialized audiences, sharing ideas and developing a team perspective as suggested by (Frank & Manning, 2007) can help researchers improve the overall output of market research which in turn can reduce bias. Extra care should be taken during interviews with participants—more conventional means should be developed when designing surveys, misleading questions should be avoided; it is imperative for the researcher to understand the strengths and weaknesses of each question type they will be using, this way the question and its' options will not only be chosen correctly but can be tailored in order to provide only the most useful data. (Penwarden, 2013).

Pretesting is an important approach to help reduce bias in market research—it is essential to use pretesting to ensure that there are no issues with the proposed choice of styling. Although pretesting can be a repetitive process, proper precision should be taken when analyzing various pretest results. Meta-Analysis is also a way to elude bias, especially in quantitative research—meta-analysis is a statistical procedure in which data collected with different techniques from different studies are combined and analyzed. Additionally, real-time experience tracking (RET) can be used to collect accurate data, while biometric belts and electroencephalography (EEG) caps are effective tools that could be used by researchers. Likewise techniques, ranging from reading facial expressions to measuring tiny differences in reaction time can also be utilized to minimize bias. Researchers can also adopt "exploring" techniques where they can spend time with consumers in their homes, place of work, and comfort zones—examining behaviors and looking for contradictions.

4 Conclusions

While it is relevant to recognize that bias is ubiquitous and must be seen predominantly as a function in a study design, not of the results, it should be addressed early in the planning and designing phase of the study. Market research has unarguably made enormous advancements both quantitatively and qualitatively. Incorporating technology overtime has helped market research significantly—market researchers and industry experts have now coined the term "Big Data" to connote the enormous amount of data available in real time. Social Media has made it relatively easy to get slightly accurate data from consumers—data mining has been a fundamental technology used by researchers to collect accurate data; data storage has become cheap and easy to access, consequently customers are naively providing researchers with vast information more than ever before.

Finally, it is important to keep in mind that cost can affect the design of a research, hence some small organizations may not necessarily adhere to some of the best practices underlined in this paper, however the future of market research remains progressive.

5 References

[1] Belle, F., & Anne, M. "Reinventing Qualitative Research." (Word Advertising Research Center), Admap. (pp. 2-4). 2007

[2] Dooley Roger. "How Warren Buffett Avoids Getting Trapped by Confirmation Bias". (Forbes.com, August 2013) November 3, 2013

[3] Dooley Roger. "The Most Common (And Dangerous) Market Research Mistake" (Forbes.com, May 2013) retrieved November 2, 2013

[4] Dooley Roger. "Why So Much Market Research Sucks" Forbes.com, (Oct 2012). Retrieved November 2, 2013

[5] Kareen Huffman. "Psychology in Action"*;* New Jersey: John Wiley and Sons, Inc (C.T. Jackson, Ed) (2010)

[6] Naresh K. Malhotra, "Bias Breakdown"; (Marketing Research A. P.), (2007).

[7] Olsen, R "Self-Selection Bias"; In P. J. Lavrakas "Encyclopedia of Survey Research Methods" (Sage Research Methods), (2008).

[8] Penwarden, R. "Avoiding Survey Bias"; from FluidSurvey.com: (August 2013).

[9] Petri, N. "4 Distortive Market Research Biases To Avoid" Sap Businsess Innovation (January 2013).

[10] Philip M, Podsakoff, S. B. "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies" Journal of Applied Psychology Vol. 88 (5), (879—903). (2003).

[11] Tortolani, R. "Introducing Bias Intentionally into Survey Techniques"; Journal of Marketing Research.(n.d)

Correlation of Employee Performance and Electronic Employee Monitoring

C. Balch
Department of Business & Administration Northern Arizona University, Yuma, Arizona, USA

Abstract - *Historically, there is a correlation between increased employee monitoring and increased employee performance. However, technology provides many employee survey methods to increase employee surveillance that were heretofore difficult, expensive, or otherwise inaccessible in ordinary business practices. Thus, modern administrators are able to quantitatively measure many employee performance metrics with relative ease. The lack of effort required for observation may decrease employee perception of administrative involvement thus decreasing the positive performance effect associated with administrative involvement.*

Keywords: Employee Surveillance, Employee performance

1 Introduction

Organizations have a vested interest in efficient and effective achievement of organizational goals. There are many theories as to how to improve employee performance. This paper reviews how these theories relate to modern employee observation tools. [1]

2 Electronic Employee Monitoring

2.1 Common Monitoring Tools

There are a variety of tools available to record employee and organizational environments. Common employee monitoring methods include:

- *Email Monitoring* - scans incoming and outgoing messages using algorithms and filters in search of content that is against organizational policies.
- *Telephone Tapping* - Record and monitor employee calls. Voice recognition may be used to flag keywords.
- *Key Loggers* - Records all keyboard activity allowing recreation of documents and forms.
- *Application monitoring* - Records time spent using various applications.
- *Customer Monitoring* - Records time spent working for clients. Frequently used in law offices.
- *Video Surveillance* - video collection of employees. Usually on organizational location or associated with organizational vehicles.
- *Location Monitoring* - tracks employees (or their vehicles) information includes routes and speed of vehicle. [2]
- *Privacy of employee data* - ability to retrieve data on employees hard drive.

2.2 Electronic Monitoring Legal Considerations

Each of these monitoring tools has legal and organizational culture ramifications. Ramifications include:

- Employee privacy,
- Need to find information when employee is not present,
- Ownership of data,
- Security and privacy of organizational data,
- Investigation of employee criminal activities,
- Considerations in use of organizational equipment for personal use and/or criminal activities,
- Review for malicious software.

2.3 Benefits of Electronic Monitoring

Electronic surveillance provides many benefits including:

- Reduction of observation costs,
- Ability to create exception reports to discover abnormal behavior requiring administrative attention,
- Continuous knowledge of employee current and historical activities and location,

3 Historical Perspectives of Monitoring and Employee Performance

The following theories describe methods to improve employee performance and provide correlation with positives and negatives of employee performance.

3.1 Scientific Management

Scientific Management [3, 4] suggests the following:

- Work can be optimized into a number of universally applicable steps that all workers and some managers/leaders follow. Scientific Management principles include:
- Replace rule-of-thumb methods with scientifically developed work methods,
- Scientifically create methods to select, train, and develop workers,
- Ensure scientifically selected workers are using the scientific work methods,
- Scientifically divide work and management.

- Some if not all employees would prefer not to work and/or "soldier" so as not to produce at their peak ability. Scientific Management seeks to identify and remove soldiering behavior. Reasons that employees soldier or encourage others to soldier include:
- Perception that greater productivity will mean fewer jobs and/or discrimination against those who are less productive.
- When paid the same, workers feel the least productive worker sets the productivity standard.
- Employees have little initiative to discover and apply best methods to do their work.

Scientific management and associated work styles clearly lends itself to effective use of electronic surveillance . Electronic surveillance allows detailed monitoring of employee actions, methods, and performance.

3.2 Hawthorne Effect

The Hawthorne experiments [5] adjusted a number employee environmental conditions in attempts to discover optimal working conditions. All adjustments resulted in improved employee. Finding included:

- Employee satisfaction as an important component to improved productivity to academic literature. Many great employers already instinctively knew/know the importance of employee satisfaction.
- Most followers are more productive and communicative when they feel their leader's positive attention/care and they feel they matter as an individual.
- There may be other confounding variables, such as the season, that changed worker productivity.
- Whenever possible help employees enjoy their work.
- Work group norms affects productivity.
- The Hawthorne effect is very close to micromanagement.
- Attention is not always desired by followers and excessive watching (micromanagement) can cause problems.
- Strengthening teams generally increases productivity.
- The work place is a social system -- happier people and people who like their workplace culture tend to be more productive.
- Positive feedback methods are not universal as not all followers respond the same needs and leaders have different reward methods and resources.
- Recent technology makes it easy for leaders to at least appear to observe their followers. Technology may not be as good at allowing leaders to demonstrate personal care and concern for their followers through personal interaction.

The Hawthorne experiments suggest that employee needs go beyond observation. While electronic observation may assist administrators, in achieving employee needs, the act of observation is not enough to insure improved employee performance.

3.3 Situational Leadership

Hersey and Blanchard [6-8] identified four administrative styles based on Task Behavior and Relationship Behavior. They argued that no leadership style is optimal for all leaders and situations. Thus, effective leaders adapt their leadership style according to the situation.

Situational Leadership Styles	
S1: Telling	Predominately one-way, directive communication from the leader to followers where the leader defines roles and the what, how, why, when and where to do the task. Often includes low supportive behavior.
S2: Selling	Higher relationship behavior includes two-way communication between leaders and followers, encourages followers to buy into the process while the leader still defines roles and direction.
S3: Participating	Increased relationship behavior includes leader fostering an environment of shared decision-making including aspects of how tasks are accomplished and assigned.
S4: Delegating	While the leader is still involved with decisions, most of the process and responsibility is delegated. The leader stays involved to monitor progress.

Situational leadership theory defines followers into four categories according to their maturity levels:

Follower Maturity Level			
Low	**Moderate**		**High**
M1	**M2**	**M3**	**M4**
Lack skills needed for the job and are not willing to take responsibility.	Lack skills but are willing to try.	Have skills but are unwilling to take on the task.	Have skills and are willing to work.

Important situational leadership considerations:

- Leadership style should match goals and followers.
- Relationships between administrator and employee should match administrator style and employee maturity level.
- Portions of the grid require the leader trust followers. Some leaders may find trust difficult.

- The leader would look for followers who fit both the work requirements and their leadership style.
- Followers are more effective when they appreciate how they are lead -- followers should feel empowered to suggest to leadership methods that best suits them.
- Both leaders and followers change depending on situation/task at hand.
- Few leaders (managers) are able to adapt to the optimal style for all situations. Many are also unable to adapt to the optimal style for their current situation but most managers find a reasonably effective style.
- Reality is far too complicated for labels to completely define.
- It is impossible for a leader to have perfect information. Leaders must settle for satisficing.

Situational Leadership suggests that employee surveillance is more effective in some situations and with some employee maturity levels. Situational theory also suggests that some employees may be discouraged by increased surveillance.

3.4 Follower Needs

Administrators should consider the fit between follower needs, performance, and surveillance. Follower needs often include:

- Pride of product/accomplishment,
- Having a voice that is heard,
- Positive feedback, acknowledgement, appreciation, and praise,
- A workplace that makes them happy,
- Knowledge that the leader and coworkers care for them,
- Personal growth,
- Belief that their thoughts matter and contribute,
- Equal opportunity,
- Camaraderie (including their relationship with leaders),
- Income,
- Prestige.

Electronic monitoring should reflect consideration of follower needs.

4 Discussion/Conclusion

While often perceived as an area of distrust, employee monitoring tools potentially improve employee performance by addressing employee needs.
Monitoring tools that help meet the needs of organizational employees administration should be embraced. [9]
For instance, monitoring needs in a call center are substantially different from monitoring needs of university faculty or CEOs. Call centers measure performance on immediate operational needs such as numbers length of customer wait and length of conversations while faculty and CEOs have much strategic concerns that allow for some immediate setbacks.
Even call center employees who do not meet immediate operational goals may be meeting strategic goals with longer conversations but greater satisfaction.
Have reasonable doubt when neatly defining persons and/or situations and the resultant need for employee monitoring. It is often difficult, if not impossible, to categorize people. Theories that attempt to categorize oversimplify and thus have varied correspondence and applicability to the real world. On the other hand, categories and models provide important considerations that assist with monitoring decisions.

References

[1] D. A. Wren, *The Evolution of Management Thought.* New York: John Wiley & Sons, .

[2] A. Shontell, "Why UPS Is So Efficient: "Our Trucks Never Turn Left"," *Business Insider,* March 24, 2011.

[3] F. W. Taylor, *The Principles of Scientific Management.* New York, NY, USA and London, UK: Harper & Brother, 1911.

[4] L. D. Parker and N. R. Lewis, "Classical management control in contemporary management and accounting: the persistence of Taylor and Fayol's world," *Accounting, Business and Financial History*, pp. 211-235, 1995.

[5] F. J. Roethlisberger and W. J. Dickson, *Management and the Worker: An Account of a Research Program Conducted by the Western Electric Company, Hawthorne Works.* Chicago: Harvard University Press, 1937.

[6] P. Hersey, *The Situational Leader.* New York, NY: Warner Books, 1985.

[7] Hersey, P. & Blanchard, K., *Situational Leadership: A Summary.* Englewood Cliffs, NJ: Prentice-Hall, 1988.

[8] P. Hersey, K. H. Blanchard and D. E. Johnson, *Management of Organizational Behavior: Leading Human Resources.* Upper Saddle River, NJ: Prentice Hall, 2001.

[9] Kroll, D., & Wang, X., "Using group technology to improve quality and response time," *Industrial Management*, pp. 21, 1994.

Integrating WeBWork in Calculus III

Larry Wang
Mathematics Department, Southern Polytechnic State University, Marietta, GA, USA

Abstract - *The aim of this paper is to demonstrate how the online homework improves the student success rates in one Calculus III course. WeBWork provided by the Mathematical Association of America is adopted as the online homework system. In the fall 2013, the system was implemented in one class of Calculus III. As a result, the passing rate of the section with WeBWork is well above other sections without WeBWork (about 7% higher). The paper also shows how the WeBWork system was used.*

Keywords: Online grading, WeBWorK, pass rate.

1 Introduction

This is a continuation of the previous project, "Improving Student Success Rate with Online Homework". After successfully implementing WeBWork as an online homework grading system in one section of PreCalculus and one section Calculus I in the spring of 2013, we have seen a tremendous increase (about 17%) in the passing rates and an improvement of grades (about 20% increase in grades of A or B). Although the sample size is small, WeBWork has had a dramatic impact on student success in those two classes. We would like to see how the system works in Calculus III classes. The system was implemented in one section of Calculus III in the fall of 2013. More functions of the system, such as statistical analysis, were used. It turns out that the system has made as significant impact in Calculus III as in PreCalculus and Calculus I. In this paper, we will demonstrate how the system was used and show the effect of the system by comparing the success rates between the classes with WeBWork and classes without WeBWork. Based on the analysis, we conclude that WeBWork as online homework system has positive impact on students learning in math courses.

2 WeBWorK

WeBWork is an open-source online homework system for math and science courses. It is supported by the Mathematical Association of America and the National Science Foundation and comes with a National Problem Library (NPL) of over 20,000 homework problems. Problems in the NPL target most lower division undergraduate math courses and some advanced courses. Supported courses include college algebra, discrete mathematics, probability and statistics, single and multivariable calculus, differential equations, linear algebra and complex analysis. WeBWork has been successfully used by over 300 colleges, universities, and high schools. Unlike other commercial online homework softwares, WeBWrok is free to students. It is also easier to manage and popularize the system. WebWrok provides the tools needed to deliver homework and grade them online. Students can do their homework from anywhere and at anytime as long as they can access the internet. The assignments do not have to be completed one time. They can also be printed out and worked by students. Students receive the feedback immediately after they submitted their answers. They can keep try until they get correct answers or reach a limit of tries set by the instructor. The instant feedback make students feel accomplished if they get right answers or make them want to redo the problems if they get wrong answers. After the due date, students can view the homework and solutions of the problems. WeBWork offers students a personalized interactive learning environment, where they can learn at their own pace and measure their progress. In addition, WeBWork provides different sets of problems for different students. This feature allows students work together, but each of them has to find solutions of homework problems independently.

The system also provides a convenient email communication between the instructor and students. While working on the homework problems, students can send the instructor an email to ask questions without retyping the complicated math symbols in other email systems.

3 Implementation

In the fall of 2013, WeBWork was adopted in one section of Calculus III offered at Southern Polytechnic State University. Students were assigned homework through WeBWork once every two weeks. Each of the assignments consists of 10-20 questions and has a due date. Usually, students have one week to finish an assignment. They were given four or five tries for each problem. The homework contributed about 15% of the final course grade. Students still take the regular quizzes, tests, and a comprehensive final exam for 85% of the course grade.

After students submitted their answers, we used the item analysis tool offered by WeBWork to check which problems students struggled with. Then we discussed those problems and reviewed the related course contents. Students really like this approach. As a result, they performed well on quizzes and exams.

Figure 1 shows a list of courses at SPSU registered to use WeBWork.

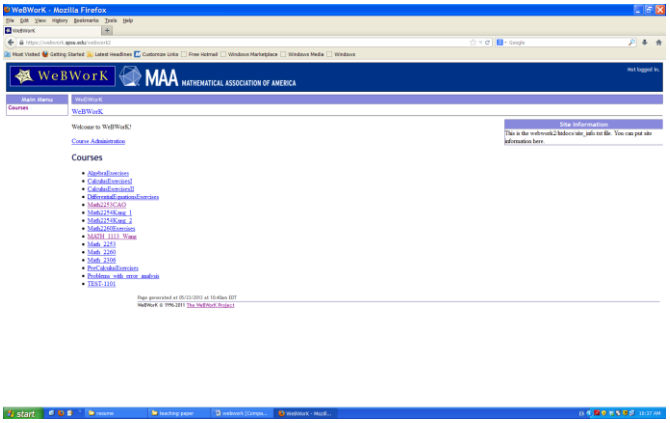


Figure 1: List of courses

Students click the course to enter the system with user name and password. Figure 2 shows a window for a student to enter his/her user name and password.

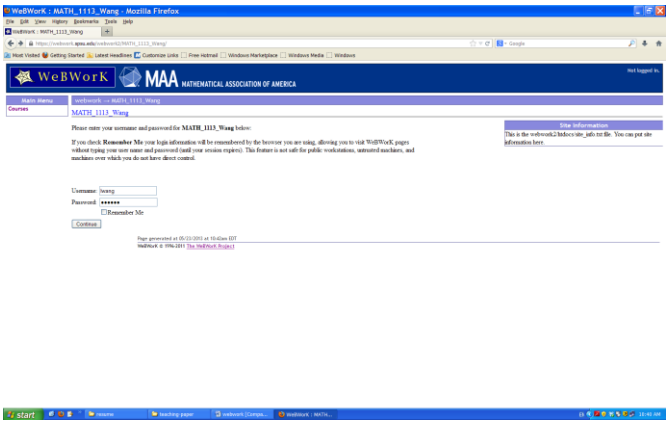


Figure 2: Login page

Once students log into the system, they will be shown a list of assignments. Figure 3 shows a window for a list of assignments.

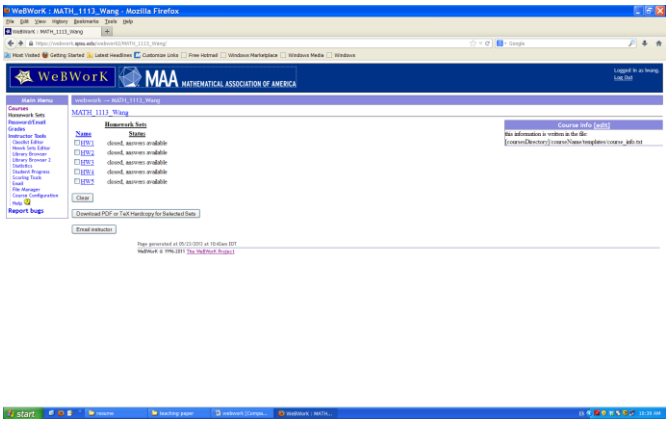


Figure 3: List of homework assignments

After students click the appropriate assignment, a collection of problems will be displayed for students to work. Figure 4 shows how problems look like.

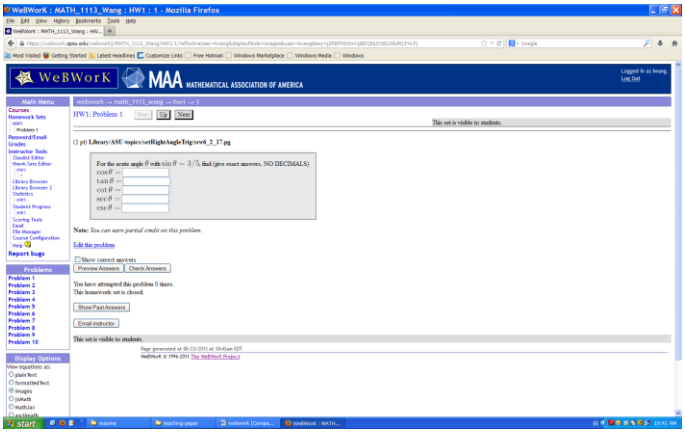


Figure 4: Problem page

After students work on the problems, they enter their answers in the boxes provided for each part of the problem.

Immediately, they can check whether the inputted answers are correct or not. They can keep trying until they reach the limit of four or five times.

4 Effects of WeBWorK

At the end of the fall of 2013, the passing rate in Calculus III enhanced by WeBWork is 7% higher than the passing rate in Calculus III classes without WeBWork. The percentage of A's and B's also increase by 12%. (See Table I).

Comparing with those sections in which WeBWork was not used, we found that not only the passing rate was significantly higher but also the retention rate increased. Here is a table to show the detailed effect:

Table I: Comparison on passing rates in Calculus III.

	A	B	C	D	F	Passing Rate
With WW	31%	36%	17%	3%	8%	84%
Without WW	24%	31%	22%	12%	5%	77%

The table shows that WeBWork improves students' performances in the class. It was found that more students were motivated to work on their homework with WeBWork. Just like students in PreCalculus and Calculus I enhanced with WeBWork, they prefer to working with computers because students in the digital age are used to learning via technology.

They like the instant response to the correctness of answer, and the option to correct the wrong answers without penalty since students need to know, as they are learning, if they are doing the problem correctly and understanding it fully. That is why most students like WeBWork as a survey, which was conducted at the end of the fall, 2013, reflects the students' feedback on the WeBWork. Here are four questions in the survey.

1. Between the traditional homework and the online homework from WeBWork, do you like WeBWork better?
2. Does WeBWork help you learn in this course?
3. Do you agree that WeBWork has more strengths than weaknesses?
4. Would you recommend others to use it?

We summarized the responses in the following table II.

Table II: Survey result

	1	2	3	4
Yes	86%	88%	75%	86%
No	14%	12%	25%	14%

5 Conclusion and Future Work

The use of WeBWork has significantly improved students' success rate in Calculus III. Both the passing rate and the students' response to survey questions support the positive impact of using WeBWork. The results consistently showed a direct correlation between required use of WeBWork for homework assignments and higher success rate.

Why did WeBWork work? The online homework system WeBWork helped make up for the lack of interaction between teachers and students. With oversized classes (usually 30 or more students in a class), the instructor does not have enough time to grade homework. Some students do not work on suggested homework problems at all. Consequently, these students do poorly on the tests. They feel frustrated and lose their interest in the subject. This is why the WeBWork played a key role in the student's learning process. It has been shown that students more likely do their homework online because many of them are used to and prefer a technological environment. By completing their homework on time, students came to class more prepared and ready to participate in discussion. Their confidences and motivations were gradually built up, which were two main factors for their successes in the course.

In the future, we would like to further analyze the impact of using WeBWork on subsequent courses. The passing rate of students taking advanced mathematics courses who came out of a redesigned Calculus III course with WeBWork will be compared with the overall passing rate of traditional courses. Other plans for the future include using even more of the tools offered by WeBWork to further increase student success rates.

6 Acknowledgements

We would like to thank students in Calculus III for their participations in the project. We also would like to thank Dr. Shangrong Deng for his help to set up the websites for the classe.

7 References

[1] Thomasenia Lott Adams. "Technology Makes a Difference in Community College Mathematics Teaching", Community College Journal of Research and Practice, 1997.

[2] E. Johnson and K. Green. "Promoting Mathematical Communication and Community via Blackboard"; Rimus: Problems, Research, and Issues in Mathematics, Undergraduate Education, Vol. XVII, No.4, October-December, 2007.

[3] C. Spitznagel. "Teaching Mathematical Writing Electronically", Proceedings of the 18th Annual International Conference on Technology in Collegiate Mathematic, March, 2006.

[4] Teresa Thiei and M. Brown. "Addressing the Crisis in College Mathematics: Designing Course for Student Success", Change: The Magazine of Higher Learning, 2008.

[5] http://WeBWorK.maa.org/intro.html.

Automating E-Portfolio Review Processes

Kevin A. Gary, Anusha Gopalakrishnan
Department of Engineering and Computer Systems
Arizona State University
kgary,agopala4@asu.edu

Abstract. **E-portfolio management systems assume very high importance to universities as a means of evidence-based assessment of student work. With the availability of lot of open source e-portfolio tools, a great deal of work can be done to customize these to very specific needs. The goal of this project is to automate an E-portfolio review process using a conference paper style collaborative review workflow. Mahara is an open source e-Portfolio and social networking web application which has been used for this purpose. This study aims at highlighting the role of technology in the growth of education.**

Index Terms— Software engineering, e-portfolio tools, Mahara, Workflow Management, Outcomes, Rubrics

I. INTRODUCTION

'An e-portfolio is a digitized collection of artifacts that includes demonstrations, resources, and accomplishments representing an individual, group, or institution' [1]. It works as an administrative tool for managing and organizing multiple sets of collective work and also controls on who can see what. E-portfolios encourage exchange of ideas and feedback among different users [1]. Arizona State University has a manual portfolio review process. Each degree program has a set of outcomes against which the department evaluates the students belonging to that program. Each student, at the end of their final year in the degree program is required to submit a portfolio for every outcome of the program. The students are required to submit enough evidence to support the outcome. The committee responsible for an outcome reviews and evaluates the respective portfolio against the outcome.

A. Problem Statement

Portfolio creation and review processes are manual processes that tax the time of both students and professors. The process involves a lot of paperwork, and storing the portfolios and assessment results and retrieving them for future reference requires physically storing the paperwork. Hence the objective of this project is to automate this portfolio review process to improve productivity and accuracy.

B. Motivation

The Department of Engineering and Computer Systems at Arizona State University describes a portfolio review process as significantly "paper pushing" [2]. Students prepare portfolio binders, one for each of the program outcomes. The binders are sent to the committee for evaluation. The evaluation results are recorded in paper and the results are provided at the end of the semester. According to an internal study [2], this 'process is laborious, time-consuming, and subject to inconsistencies'. The department has to maintain all the binders and the paper evaluations for all students which in turn is space consuming and increase with the increase in students' intake. An e-Portfolio platform has the potential to speed the process and reduce clutter. By having an electronic record, faculty can have a formal review process by maintaining records of all communication with the students and within the committee. The system has the potential to serve as a data repository and can help in other administrative tasks like performance evaluation at program level or

at a much higher departmental level. Creating portfolios electronically will be much easier for students than preparing paper portfolio binders. It is easy to assemble portfolios per outcome basis. Updating the portfolios based on formative feedback is also much easier electronically when compared to the manual process. Students can maintain these portfolios and the evidence for future reference. They have the advantage of sharing their work for future potential employers or graduate schools

Meyer and Latham [3] conducted a research on the benefits and challenges of implementing e-portfolios. They conducted a survey among four universities that implemented a common portfolio system and recorded their experiences with the tool. Using this system the students can create lesson plans, portfolios, and other projects. These items created by students are assessed by instructors using a set of rubrics. Some of the responses during the survey are 'The tool is an efficient means of managing and analyzing data, particularly assessment data, are a tremendous benefit' [3], '[The e-folio system] encouraged common requirements across all teacher education programs.'[3].

Thus using an automated tool for carrying out the portfolio review process has significant advantages over slow and tedious manual process. It is quite clear that this tool can act as an effective data management tool which, to a great extent can reduce the manual efforts of storing and retrieving paper based portfolios of students for evaluation purposes.

C. Applications

An e-portfolio tool is not just a data management tool; it does more than that.

1. Accreditation and Departmental Review: This tool can also be viewed as a tracking system that records the performances of students. These archives on students' work collected over a large period of time truly depict the performance growth of a department or an institution and hence can serve as a strong evidence for departmental or institutional accreditation. The system can also play an important role in inter departmental review process by comparing the outcomes evaluation results of various programs within the same department[4]
2. Academic Progress: The e-portfolio tool helps in creating a structured workflow between a student and a committee. The tool makes it easy for the committee to review and keep track of all past and present activities of a student. This in turn makes it very easy for the department to keep track of a student's progress by pulling out periodic reports on the student from the tool [4].
3. Career Opportunities: E-portfolios can be used by students even after their degree completion. The tool allows students to save their academic work, feedback from faculty members. Students can create a whole new portfolio from their previous academic work which they can present to a potential employer. In this way the employers not only get a chance to see the progress of the students through the supporting evidence but can also look at the feedback of a faculty and other academic people on the student's work [4].

II. BACKGROUND

This section lists some of the available open source portfolio systems and explains why Mahara was chosen as the development platform.

The Open Source Portfolio (OSP) is an online repository where the users can store their work. It also serves as an environment in which a user is able to exhibit their work in various forms by creating different views or portfolios. 'A portfolio owner can collect items that best represent their accomplishments, their learning, or their work; to reflect upon these items and their connections; to design a portfolio that showcases the best selections of this work; and to publish the portfolio to designated audience' [5].

'Common Interest Groups (CIG)' users [5] who are part of the open source portfolio development communities and administrators assist the portfolio owners in portfolio development. The portfolio owners can give access to other users who can be CIG coordinators, evaluators, reviewers or guest users to view and review their portfolios and provide formal evaluation or informal feedback [5].

There are many such tools available and this section will attempt to briefly describe some of the popular tools.

1 Mahara

Mahara is an open source ePortfolio and social networking web application. Mahara is an online repository for storing important data in any format like documents, videos, pictures, code. It allows users to create and maintain a digital portfolio by putting together relevant information in a structured format as preferred by the user. The social networking features of Mahara allow users to interact with each other. Mahara provides users with blogs, a resume builder, a file manager and a view creator - a tool to help users create arrangements of their content in a particular way for others to see [6].

2 Open Source Portfolio – Sakai Project

Sakai is a 'Collaboration and Learning Environment (CLE)' [7] which is an open source, educational software platform that is used for teaching, research and collaboration. Sakai provides various functions like document distribution, gradebook, chat tools, discussion forums, online test features, assignment uploads that are common features of a course management system. Sakai project has an additional tool called 'open source portfolio'. This tool, similar to other portfolio tools provides an environment for users to create portfolios and exhibit their work [7].

Mahara appears to be a good candidate to perform the portfolio review automation process. Mahara has existing capabilities of creating different views and upload materials like documents, videos, blogs etc. It allows the users to share the views with other users and get feedback on their views. This functionality can be extended as the formative feedback method which is a part of the portfolio review process.

III. IMPLEMENTATION

The Mahara Open Source portfolio system was used to build the e-portfolio tool. Detailed workflow, database design and implementation process are discussed in this section.

A. E-Portfolio Requirements

1. Automated portfolio review process with student, committee and chair playing different roles in the process.
2. Creation and maintenance of set of outcomes specific to a degree program.
   Example: Software Engineering program and Engineering program has relatively similar set of outcomes. Using this tool two set of outcomes can be created one for each program. The students will be evaluated only against the outcomes for the program they belong to.
3. Tool should provide at least two different assessment options for an outcome. *Level based assessment* – Students are assigned levels (from 1 to 4 or more) against each outcome or *Yes/No based assessment* – Student either meets the requirements for the outcome or does not meet.
4. Tool should support complex outcome structure: It should allow the creation of sub outcomes for the outcomes and further more sub outcomes for those sub outcomes and so on. The portfolio submitted to the main outcome committee should automatically be available for all sub outcomes committees.
5. The tool should support access based outcome evaluation: Certain outcomes to be created for specific focus areas within a program. Thus only those students belonging to a specific primary focus area can be evaluated against the outcomes for that focus area.
6. Rubrics for outcomes: Tool allows the creation of set of rubrics for each outcome. Then the portfolio is not only evaluated against the outcome but also against each rubric and final result is consolidation of all rubrics' result.

B. Design

The design section describes the portfolio review process. The process is described through a sequence of steps as seen by or implemented in the system and hence ties the review process directly to the workflow injected into the Mahara system. Second part of the design talks about the changes made to the database to incorporate the portfolio review process without affecting the existing system, thus effectively maintaining the core functionalities intact.

1. Portfolio Review Process

This section describes the portfolio review process through UML sequence diagrams.

Step 1: Admin task – creating outcomes

System administrator creates set of outcomes for each degree program. He sets the assessment type to be used and adds rubrics, if available. He then creates template portfolios one for each outcome and makes it available for the students. He also creates committees, one for each outcome and adds members with different roles such as student, committee member and chair. All students of a degree program are members of all committees with the role as student. (Figure 1)

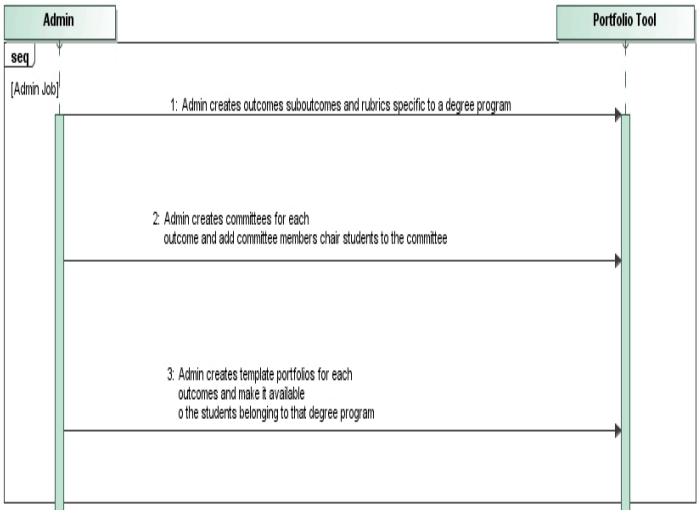


Figure 1: Outcomes and Committee creation

Step 2: Formative Assessment

Students start creating portfolios for the outcomes from the templates. If, for example, there are eight set of outcomes in a degree program, then each student has to create eight different portfolios one for each of the eight outcomes. The students can request the respective committee to review their portfolios and provide a formative feedback based on which they can improve their portfolios. (Figure 2)

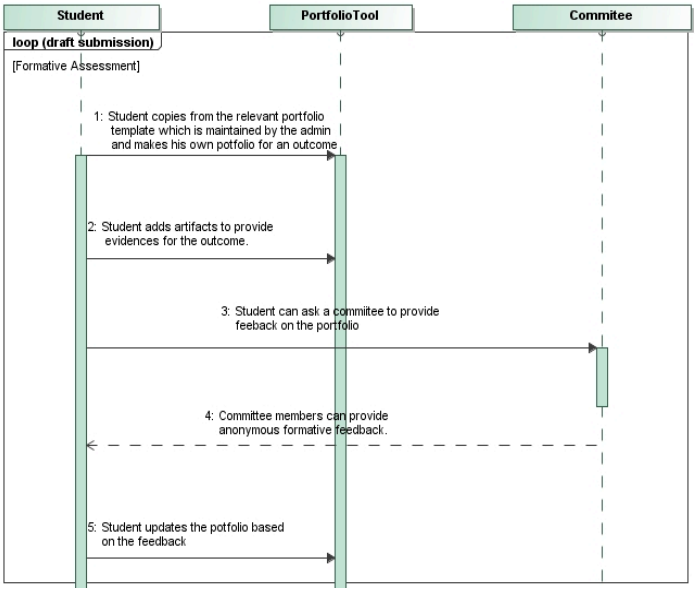


Figure 2: Formative Assessment

Step 3: Summative Assessment - Student

Once the portfolios are ready, the student submits each portfolio to the respective committee. Before submission, the student does a self evaluation of his portfolio against the outcome. Once submitted, the student will not be able to edit the portfolio anymore. (Figure 3)

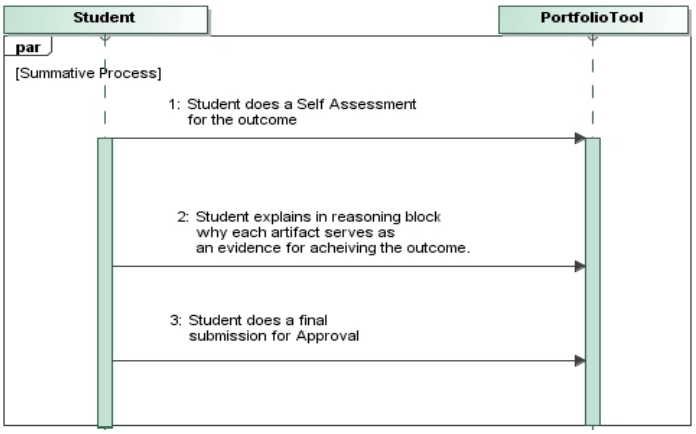


Figure 3: Summative Assessment - Student

Step 4: Summative Assessment - Committee

Once the portfolios are submitted by the student, the committee members will be able to view all the portfolios submitted to their committee. Each committee member individually evaluates the portfolio against the outcome and rubrics. They then will submit their assessment to the chair of the committee. (Figure 4)

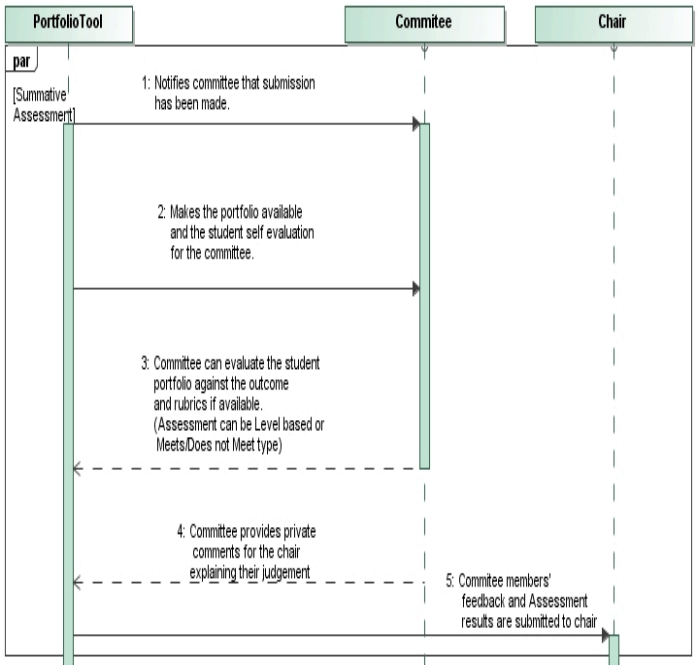


Figure 4: Summative Assessment - Committee

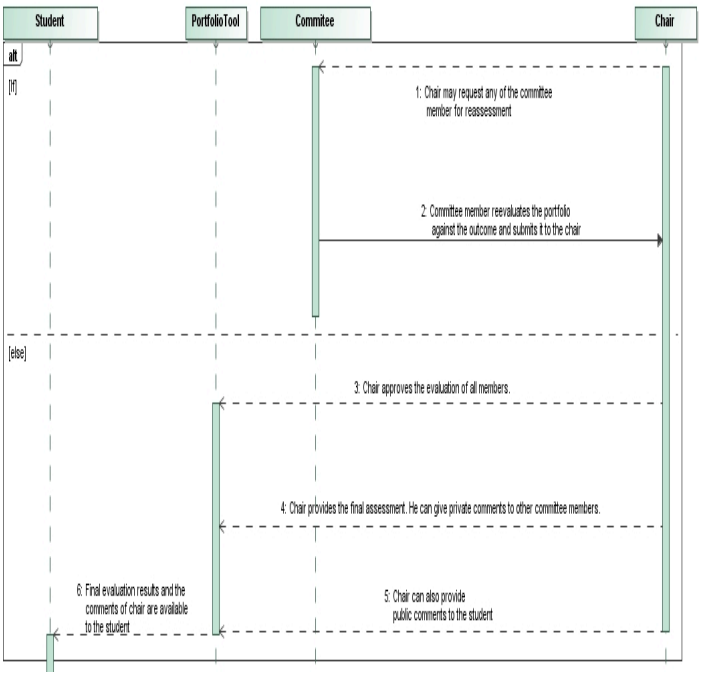
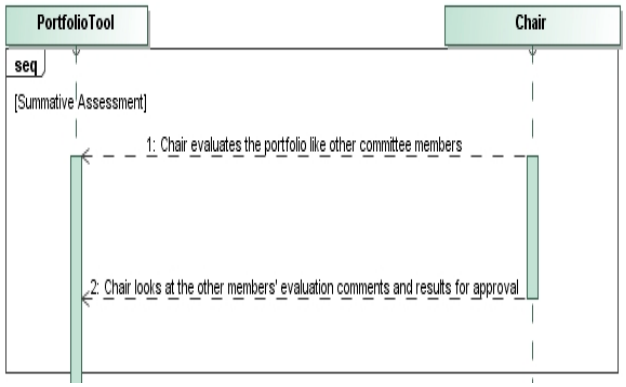


Figure 5: Summative Assessment - Chair

Step 5: Summative Assessment - Chair

A chair is the head of the committee who is in charge of evaluating all the assessment results submitted to him by other committee members and post the final results to the student. Once the committee has done assessing the portfolios, the chair will then be able to view a report on the assessment of every student's portfolio by the all committee members. Chair can either approve their assessment or ask a committee member to reassess one or more portfolios. (Figure 5)



2. Sub Outcome – Main Outcome Review

An outcome can be further divided into sub outcomes and the sub outcomes can have committees different from the main outcome committee. Then the portfolio which is submitted to the main outcome committee is automatically submitted to all sub outcomes committee. The sub committees assess the portfolio against the outcome they are in charge of. Only when all the sub committees are done with the evaluation, the main committee can start its assessment. (Figure 6)

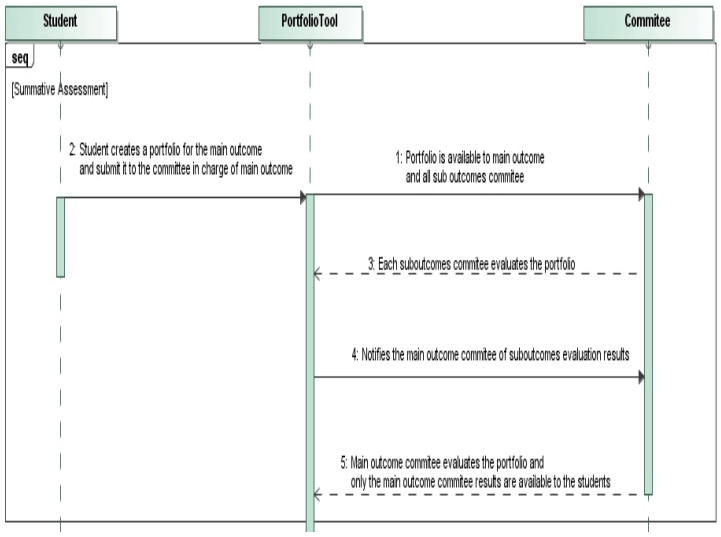


Figure 6: Sub outcomes – Main outcome workflow

C. Technical Details

This section explains the technical implementation details. Similar to database design changes, this section explains the code level changes made to the existing Mahara code base to implement the workflow and ensuring that the existing functionalities are not affected.

1. System Details

Code Base: Mahara (version 1.2.2)
Database: MySQL database
(mahara database which is a part of Mahara System)
Programming Language: PHP

2. Coding Details

Portfolio review workflow is built around the existing functionalities of Mahara. This involved coding in PHP language. Existing Mahara system has an option of creating groups and adding users to the group with different roles. This functionality was extended in the process such that the groups serve as committees. Three roles such as student, committee member, and chair were created and users were added to the committee with appropriate roles. New screens were added to create outcomes and link it to a committee. Users can create portfolios using existing Mahara system. This functionality was modified to track each portfolio to an outcome and to allow students to submit the portfolio created for an outcome to the corresponding committee in charge. New screens were added for the evaluation process.

17 new php source code files were created and added to the existing Mahara source code. These files cover the major functionalities such as creating outcomes, committee and chair evaluation process, and sub outcomes valuation. 16 existing Mahara source files were updated to merge the workflow process as a part of the existing Mahara system.

3. Database Implementation

Mahara has a built in database that stores users' information, portfolios or views, artifacts added to each portfolio, committees or groups, details of members, roles within a group, portfolios submitted to a group. Database design includes creation of new tables for storing outcomes, rubrics information, assessment results of all portfolios against outcomes and rubrics. Existing tables were modified to interact with the new tables. Careful design changes were made to Mahara database such that the modified tables work along with the existing database.

New tables created include:

- *outcomes* – Stores all outcomes for all degree programs. This allows a suboutcomes structure by having a referential integrity constraint to itself.
- *outcome_levels* – This table is used to add Level-based assessment to the outcome
- *rubrics* – storeS rubrics specific to an outcome.
- *outcome_results* – Stores all evaluation results of all students against all outcomes and rubrics.
- *degree_programs*, *primary_focus* – Tables for storing degree programs and various specialization available for each.

The diagram below shows attributes and referential constraints between some of the new tables.

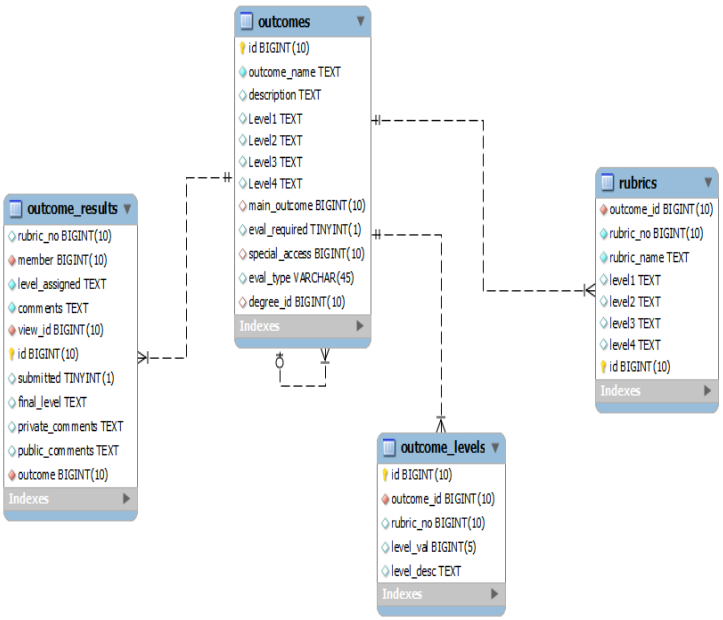


Figure 7: New tables add to Mahara schema

V. CONCLUSION AND FUTURE WORK

The e-portfolio tool has been used in 3 successive iterations of a junior software engineering project course at ASU. No formal evaluation of the utility of the system was conducted, but there are anecdotal stories, good and bad, of its application. On the positive side, the tool avoids the paper-based process previously used, and gives the students an online convenient mechanism to upload evidence. On the

negative side, it is clear an e-portfolio system such as this still does not address fundamental difficulties with evidence and reflection-oriented assessment. It is very difficult to conduct committee-style reviews in a timely fashion even with the convenience of an e-portfolio system. Finally, student course evaluation comments indicate that while the e-portfolio system itself seemed to be generally accepted, students questioned why it was distinct from the course management system and did not like to do the extra work of accumulating and reflecting on evidence.

A database driven e-portfolio tool has been developed in this work that meets the predefined portfolio review process requirements. Also, a set of database tables that stores the outcomes information, sub outcomes their mappings and all portfolios outcomes evaluation results are designed and implemented. The user friendly tool developed in this work provides dynamic information about the outcomes and their rubrics for each degree program.

The e-portfolio application employs many of the best practices that are generally recommended while developing web applications. It addresses separation of concerns by having separate code for implementing core functionalities, creating front end pages and for database operations. The application also follows the coding methodologies used by the Mahara code group. It uses transaction for database updates in order to maintain consistency of database and properly closes the connections to prevent wastage of system resources.

The tool can be integrated with the Arizona State University single sign on system, thus extending its use to all departments at ASU. Since, extending the tool to be used by all departments will increase the number of potential users on the system, this process requires exploring ways to make the tool withstand high traffic flow during specific time periods like end of semesters when all portfolios have to be reviewed within a deadline.

The tool can be extended to support additional reporting capabilities that can pull all the past data, consolidate and produce various useful reports such as annual reports on students for each degree program, overall report for the department with many degree programs, four year report for each program, best and worst performances of students per semester, per year or during four year period.

REFERENCES

[1] Lorenzo, G, & Ittelson, J. (2005a, July). An overview of e-portfolios. *EduCause Learning Initiative Paper 1*: *2005*

[2] Internal Document – The Version 2.0 Committee, Department of Engineering, Arizona State University, Polytechnic Campus (May 2010)

[3] Meyer, B and Latham, N. (2008), Implementing Electronic Portfolios: Benefits, Challenges and Suggestions, EDUCAUSE Quarterly, Vol. 31 (Jan-March), No.1, pp.34-41.

[4] Reese, M., & Levy, R. (2009). Assessing the future: E-portfolio trends, uses, and options in higher education (Research Bulletin, Issue 4). Boulder, CO: EDUCAUSE Center for Applied Research

[5] Open Source Portfolio: http://confluence.sakaiproject.org/pages/viewpage.action?pageId=22970395

[6] Mahara: http://mahara.org/about

[7] Sakai: http://sakaiproject.org/

[8] Sampath Kumaran, V. Web-based Course Information System supporting Accreditation using ASP.NET', Division of Computing Studies, Arizona State University

[9] Nagappan, Y. 'Workflow Management for Software Enterprise Pedagogy', Department of Engineering, Arizona State University