Digital Publishing and Mobile Technologies

15th International Conference on Electronic Publishing,
June 22-24, 2011, İstanbul, Turkey

Proceedings

Ankara, 2011

Digital Publishing and Mobile Technologies

15th International Conference on Electronic Publishing,
June 22-24, 2011, İstanbul, Turkey

Proceedings

Editors:
Yaşar Tonta, Umut Al, Phyllis Lepon Erdoğan, Ana Alice Baptista

Hacettepe University
Department of Information Management
Ankara, 2011

Digital Publishing and Mobile Technologies: 15th International Conference on Electronic Publishing, June 22-24, 2011, İstanbul, Turkey, Proceedings

http://www.elpub.net/

Publisher: Hacettepe University Department of Information Management

http://www.bby.hacettepe.edu.tr

ISBN 978-975-491-320-0

International Conference on Electronic Publishing (15th: 2011: İstanbul)
Digital publishing and mobile technologies: proceedings / 15th International Conference on Electronic Publishing, June 22-24, 2011, İstanbul, Turkey; Ed.: Yaşar Tonta, Umut Al, Phyllis Lepon Erdoğan, Ana Alice Baptista.-- Ankara: Hacettepe University Department of Information Management, 2011.
x, 171 p.
Includes bibliographic references and index.
ISBN 978-975-491-320-0

1. Electronic Publishing – Congresses I. Tonta, Yaşar. II. Al, Umut. III. Erdoğan, Phyllis Lepon. IV. Baptista, Ana Alice.
Z286.E43 In819 2011
070.5 In819 2011

General Chair
Yaşar Tonta (Hacettepe University, Turkey)

Program Chair
Ana Alice Baptista (University of Minho, Portugal)

Program Committee

Ana Maria Ramalho Correia, (UNL, Portugal)
António Dias de Figueiredo (University of Coimbra, Portugal)
Arun Iyengar (IBM Research, USA)
Bob Martens (Vienna University of Technology, Austria)
Carlos Sousa Pinto (University of Minho, Portugal)
Carol Tenopir (University of Tennessee, USA)
Christina Lioma (Konstanz University, Germany)
Emma Tonkin (UKOLN, UK)
Eva Mendéz (Universidad Carlos III, Madrid, Spain)
Fazlı Can (Bilkent University, Turkey)
Fernanda Ribeiro (Porto University, Portugal)
Gabriel David (University of Porto, Portugal)
Georg Güntner (Salzburg Research, Austria)
Harald Krottmaier (Graz University of Technology, Austria)
Heather Morrison (BC Electronic Library Network, Canada)
Hélio Kuramoto (IBICT, Brazil)
Ian M. Johnson (The Robert Gordon University, Scotland)
İrem Soydal (Hacettepe University, Turkey)
Jaime Delgado (Universitat Politècnica de Catalunya, Spain)
Jan Engelen (Katholieke Universiteit Leuven, Belgium)
Jeppe Nicolaisen (Royal School of Library and Information Science, Denmark)
Joaquim Arnaldo Martins (University of Aveiro, Portugal)
John Gathegi (University of South Florida, USA)
John Smith (University of Kent at Canterbury, UK)
Jordan M. Scepanski (Jordan Wells Associates, USA)
Jos van Helvoort (The Hague University of Applied Sciences, The Netherlands)
José Borbinha (INESC-ID / IST – Lisbon Technical University, Portugal)
Jose Carlos Teixera (Mediaprimer, Portugal)
Julie Griffin Speer (Georgia Institute of Technology, USA)
Karel Jezek (University of West Bohemia in Pilsen, Czech Republic)
Leslie Chan (University of Toronto Scarborough, Canada)
Lisa Lena Opas-Hänninen (University of Oulu, Finland)
Maria Francisca Abad Garcia (Valencia University, Spain)
Maria Nisheva-Pavlova (Sofia University, Bulgaria)
Markus Schranz (Pressetext Austria, Austria)
Micheal Mac An Airchinnigh (Trinity College Dublin, Ireland)
Milena Dobreva (University of Strathclyde, Scotland)
Nikola Ikonomov (Institute for Bulgarian Language, Bulgaria)
Orçun Madran (Atılım University, Turkey)
Paola Gargiulo (CASPUR, Italy)
Paulo Quaresma (University of Évora, Portugal)
Peter Linde (Blekinge Institute of Technology, Sweden)
Sely M.S. Costa (University of Brasilia, Brazil)
Serap Kurbanoglu (Hacettepe University, Turkey)
Silvana Gregório Vidotti (UNESP, Brazil)
Sirje Virkus (Tallinn University, Estonia)
Stefan Gradmann (University of Hamburg, Germany)
Sueli Mara Ferreira (São Paulo University, Brazil)
Susanna Mornati (CILEA, Italy)
Turid Hedlund (Swedish School of Economics and BA, Helsinki, Finland)
Umut Al (Hacettepe University, Turkey)
Wolfram Horstmann (University of Bielefeld, Germany)
Yurdagül Ünal (Hacettepe University, Turkey)

Local Committee

Ertuğrul Çimen (Kadir Has University, Turkey)
İrem Soydal (Hacettepe University, Turkey)
Yurdagül Ünal (Hacettepe University, Turkey)
Abdullah Turan (İstanbul Bilgi University, Turkey)
Didem Önal (Kadir Has University, Turkey)
Elif Cankurt (Kadir Has University, Turkey)
Eyüp Ulugöl (Kadir Has University, Turkey)
Güleda Düzyol (Hacettepe University, Turkey)
Haydar Yalçın (Hacettepe University, Turkey)
Metin Şahin (Kadir Has University, Turkey)
Nursen Ulusoy (Kadir Has University, Turkey)
Ramazan Çelik (Kadir Has University, Turkey)
Sami Çukadar (İstanbul Bilgi University, Turkey)
Tolga Çakmak (Hacettepe University, Turkey)
Zehra Taşkın (Hacettepe University, Turkey)

Symposium Web Site Design and Programming

R. Orçun Madran

Cover, Poster and Graphic Design

N. Erol Olcay

Sponsors

Elsevier

Hacettepe University

Kadir Has University

Emerald

Mikro Bilgi Kayıt ve Dağıtım A.Ş.

Springer

Informascope

EBSCO Information Services

Pearson

Reeder

Swets

Hiperlink

Anatolian University Libraries Consortium

Turkish Librarians' Association

Association of University and Research Librarians

Association of School Librarians, Turkey

Preface

Organized by the Department of Information Management of Hacettepe University and hosted by Kadir Has University, the "15th International Conference on Electronic Publishing" (ELPUB2011) took place in İstanbul, Turkey, from 22-24 June 2011. The theme of the conference was “Digital Publishing and Mobile Technologies”. This was a most timely theme: the Horizon Report 2011 identified both electronic books (e-books) and mobiles as technologies to watch as they will likely shape our ways of working, learning, studying and entertainment in the immediate future (i.e., within the next 12 months) (http://net.educause.edu/ir/library/pdf/HR2011.pdf). Mobile devices (cell phones, iPads, and tablet PCs, among others) and e-books will also have a tremendous impact on how we communicate and collaborate with people, and access, share and generate information.

Some 30 papers were submitted to ELPUB2011. All papers were subjected to a double-blind reviewing process and 20 full papers, and 4 workshop proposals were selected to be included in this proceedings book. Accepted papers come from 14 different countries (Austria, Bulgaria, Denmark, Germany, England, India, Iran, Ireland, Italy, Malaysia, The Netherlands, Spain, Turkey, and USA) and address a number of issues dealing with, mobile access to digital cultural heritage, intelligent digital publishing software, mobile learning, institutional repositories, digital publishing tools, open access, and personal information management.

We would like to take this opportunity to thank our conference plenary keynote speaker Professor Ayşe Göker (City University London, UK), and invited keynote speakers Dr. Zeynep İnanoğlu (Google Turkey, Turkey) and Professor Fazlı Can (Bilkent University, Turkey). Professor Göker talked about the digital publishing in the mobile environment and discussed the distinctive nature of mobile search and retrieval including the context in which mobile users are engaged (e.g., task in hand, environment, device, individual interests and social aspects). Dr. İnanoğlu’s speech focused the Google Art Project (GAP) and she discussed GAP’s role in making works of art more accessible and in creating a new generation of art lovers who can experience, enjoy and appreciate art works using state of the art technology including mobiles. Professor Can outlined the major developments in information retrieval within the last 70 years and concentrated on mobile information retrieval along with the opportunities and challenges wrought about by mobile devices. Several papers and workshops presented during the conference expanded the main topics outlined by our keynote speakers. We thank our authors, workshop presenters and contributors to the “pecha kucha” session on digital publishing in different countries (Belgium, Brazil, Canada, Mozambique, Portugal, Spain, Sweden, and Turkey).

Our thanks go to our host Mr. Ertuğrul Çimen, Kadir Has University’s Library Director, and his colleagues, who spent countless hours to handle local arrangements and did an impeccable job. We also thank our colleagues Mr. Orçun Madran and Dr. Erol Olcay for designing the conference web site as well as the art work. Last but not the least, it is a pleasure to thank our sponsors whose names and logos are listed in the preliminary pages of the proceedings book.

Yaşar Tonta
General Chair, ELPUB2011
Department of Information Management
Hacettepe University
Ankara, Turkey

Table of Contents

Information in Context: The Mobile Environment

Ayşe Göker

Centre for Interactive Systems Research, Department of Information Science, School of Informatics, City University London. Email: ayse.goker.1@soi.city.ac.uk

Abstract

People are ever busier, and increasingly want useful information in easily digested bite size pieces, delivered to them as efficiently as possible. The use of mobile devices combined with current and future information access and retrieval technologies can rejuvenate the publishers' existing offerings as well as suggesting new ones. However, publishers cannot simply cut and paste content to fit the display constraints of a small device, and the opportunities of the mobile environment extend beyond the use of location.

Individuals use mobiles in many aspects of their lives including work and leisure, and publishers need to be aware of the different orientations of users depending on their context of use. Mobile search has a distinctive nature which is different from traditional desktop-based searching. The success of digital publishing in the mobile environment will depend not only on the design and presentation of the underlying content, but also on nature and effectiveness of the mobile search facilities. Furthermore, in this more dynamic environment, there are many more circumstances of use and shifts in context, driven by links to the physical world and triggers within it.

Information is a key part of our lives. However, the amount of available digital information continues to grow at a tremendous rate along with increasingly diverse forms of media and communication channels. To mitigate the effects of information overload, we need to create paths through the information space for users to navigate and manage their needs. The key enabler for this is to use context information. Context information provides an important basis for identifying and understanding people's information needs. A key challenge is making more information accessible whilst also ensuring it is relevant and useful for users' information needs.

Context includes aspects of the situation, such as location, but can also include the user's task, their environment, the device that they are using for accessing information, their personal interests, and their social interactions. Additional reasons for the importance of context include: timely delivery, better matching of user expectations and experience, and better potential for linking with advertising. This was evident in early work on personalization of web search and is increasingly clear for the mobile information environment.

User studies are essential for designing and evaluating new products and methodologies that meet the needs of real users. It is important to test developed applications in naturalistic contexts and not to make only theoretical assumptions about users' needs and activities. This presentation will argue that user studies should be conducted in a realistic way and will provide example applications from travel and tourism.

The future of electronic media depends on refining our understanding of what constitutes the step-change in mobile usage and developing innovative applications to satisfy emerging needs.

Keywords: *Mobile search; mobile information retrieval; contextual information.*

Short Bio

Dr. Ayşe Göker is a senior academic at City University London. Her research since the early '90s has focused on developing novel search techniques and environments, with an emphasis on personalized and context-sensitive information retrieval and management systems. These occur particularly within mobile and wireless computing, and also in bibliographic and web environments. Her skills are in identifying user needs and developing innovative systems that meet them. On the teaching

side, Ayşe has developed course modules in information systems at both postgraduate and undergraduate levels.

She has been successful on international collaborations, with designing innovative projects and managing teams to implement them. She co-proposed and was project leader for the AmbieSense EU-IST project, whilst at Robert Gordon University, Aberdeen, Scotland. AmbieSense was a large project on ambient, personalised and context-sensitive information for travelers and tourists. Ayşe followed the project through as co-founder to establish AmbieSense as a company. Other research projects have been in adaptive information systems, image retrieval, and contextual information retrieval.

She has a BSc Honours in Computer Science and a PhD in Information Science from City University London. After her PhD she went as an academic to Bilkent University, Ankara, Turkey. Later she obtained a readership in Aberdeen**,** Scotland before returning to City. She has published widely in information retrieval, adaptive systems, machine learning, mobile and pervasive/ubiquitous computing, and interactive systems. She is on the Editorial Board of *JASIST*, and has recently edited a book in *Information Retrieval: Searching in the 21st Century*, by Wiley, 2009. She holds a lifetime Enterprise Fellowship from the Royal Society of Edinburgh and Scottish Enterprise. More recently she was selected for the Massachusetts Institute of Technology (MIT) Entrepreneurship Development Program in Boston, USA. In her profession, she has been the Chair of the British Computer Society's Specialist Group in Information Retrieval, BCS IRSG (2000-2005). She became a finalist in the Blackberry Women & Technology Awards (2005) for Best woman in technology (Academia).

Google Art Project: Democratizing Art

Zeynep İnanoğlu

Product Marketing Manager, Google Turkey, İstanbul, Turkey. Email: zeynepi@google.com

Abstract

From the teenager browsing the web in an Internet cafe in Istanbul, to the student in New York seeking inspiration for a dissertation, the Google Art Project aspires to connect more people to art. Initially developed by a group of Google employees as a side project, the project aims to make art more accessible to the masses through state of the art technology, and to remove physical barriers between the individual and the artwork. Far from preventing people from visiting the museums in person, this unrivalled access to art will encourage the Internet generation to interact with art in new ways and ultimately inspire them to visit the 'real thing'. Starting with 17 museums, the Google Art Project aspires in the near future to create a variety of digital spaces where the new generation art lovers and art institutions will be brought together seamlessly.

Keywords: *Google Art Project; art museums; digital museums.*

Short Bio

Dr. Zeynep İnanoğlu is currently Product Marketing Manager at Google Turkey, responsible for all consumer marketing operations. Zeynep holds an undergraduate degree in Electrical Engineering from Harvard University and a Master's and PhD degrees in Information Engineering from University of Cambridge. Her PhD thesis focuses on the area of voice conversion in speech signal processing and has recently been awarded the "2011 Best Paper Award" by EURASIP (European Association for Signal Processing). Zeynep has worked in various start-up ventures, including Angel.com and Phonetic Arts and has always believed in a balance of academic research and entrepreneurial pursuits.

Mobile Information Retrieval: Memex on the Move – Nothing can be more Surprising than Life

Fazlı Can

Bilkent University, Bilkent Information Retrieval Group, Computer Engineering Department, Bilkent, Ankara 06800, Turkey. Email: canf@cs.bilkent.edu.tr

Abstract

The age of information retrieval (IR) system research is almost the same as that of digital computers. In 1945, a hypothetical hypertext-based personalized IR machine called Memex was described by Vannevar Bush in his famous *Atlantic Monthly* article. The Memex, a microfilm-based theoretical device, was designed to store and search all books, communications, and records a person had accumulated. During their long lives, IR systems have achieved great progress; they are much more surprising than the Memex. This progress comes from advances in computer science and engineering (CSE), information technologies (ITs), and innovations that happen by synergistic interaction of these fields.

Discoveries in IR have not been as significant as those of the natural sciences. But on the other hand, IR studies, the efforts of many people, together with the advances in CSE and IT have drastically changed our lives. Change has occurred especially in the last two decades with the development of Web search engines, browsers, and related applications.

IR systems were first based on mainframe computers and only accessible by a few number of specialists. After the invention of multi-user operating systems and terminal technology, they became accessible by many professional people at the same time. Then came inexpensive mass storage technologies, powerful PCs, the Internet and Web technologies; now ordinary people at any age can access almost any information (and perhaps too much information) from anywhere with a great speed. Now people feel the pressure of information glut and still cheerfully access information from their mobile devices. In other words, information that people need and don't need is now as mobile as people themselves. Mobile IR systems bring many great opportunities to its users and even greater challenges to its researchers.

Mobile IR research area overlaps with various research areas in CSE and ITs. Among others, these include augmented reality applications, computer networks, human-computer interfaces, information filtering, location-based services, machine translation, novelty detection, personal information management, and recommendation systems. Too much information, and the possibility of accessing one's own knowledge base from anywhere at any time, bring the capabilities of current mobile information systems beyond the imagination of its builders. In this talk, I will consider the changes in IR systems by considering the questions of how, when, and why those changes occur? Emphasis will be given to mobile IR.

Keywords: *Historical perspective; information glut; information retrieval; Memex; mobile information retrieval, Web.*

Short Bio

Fazlı Can received his BS degree in Electrical Engineering and his MS and PhD degrees in Computer Engineering from the Middle East Technical University, Ankara, Turkey. During his PhD studies in 1982 and 1983, he was a research assistant at Arizona State University and worked as an engineer at Intel Corporation in Phoenix, Arizona on the RAP database machine project. He is a professor of Computer Engineering at Bilkent University, Ankara, Turkey. Before joining Bilkent University he was a tenured full professor at Miami University, Oxford, OH. He was one of the two co-editors of the

ACM SIGIR Forum between 1995 and 2002. He has published several papers in computer science conferences and journals and received the Sigma Xi Miami University Researcher of the Year Award in 1998. He has served on the program committees of several international conferences, workshops, and NSF and TÜBİTAK panels. His more recent funded research is on new event detection and topic tracking in web portals, novelty detection in topic tracking, and analysis of Ottoman texts and poetry. He is one of two co-founders of the Bilkent Information Retrieval Group.

A Mobile Scenario for Electronic Publishing based on the MIPAMS Architecture[1]

Jaime Delgado, Silvia Llorente, Eva Rodríguez *and* **Víctor Torres-Padrosa**

Departament d'Arquitectura de Computadors, Universitat Politècnica de Catalunya, C/Jordi Girona, 1-3, 08034 Barcelona, Spain. Email: {jaime.delgado, silviall, evar, vtorres}@ac.upc.edu

Abstract: *This paper describes several scenarios for the management of digital media, focusing on electronic publishing from mobile environments. The solution proposed in those scenarios is based on MIPAMS (Multimedia Information Protection And Management System), a service-oriented Digital Rights Management (DRM) platform, which enables the creation, registration and distribution of multimedia content in a secure way, respecting intellectual property rights. The particularity of the mobile scenario with respect to others is the limited capability of mobile devices. A specific use case has been identified for the mobile environment and a new system, based on MIPAMS, has been designed for the electronic publishing environment.*

Keywords: *Content management; mobile environments; digital rights management.*

Introduction

This paper proposes a solution for the management of digital media in mobile environments. This solution is based on MIPAMS (Multimedia Information Protection And Management System) (Delgado, 2011), a service-oriented digital rights management (DRM) platform developed by the authors, which enables the creation, registration and distribution of multimedia content in a secure way, respecting the intellectual property rights. The proposed solution has been designed for mobile devices, which have limited capabilities, and applied to an electronic publishing scenario.

The paper is organized as follows: First, a set of content management scenarios is presented. In all of them DRM architectures facilitate the development of alternative applications. They include DRM-enabled content access control, content licensing and intellectual property registry with external licensing. Then, the MIPAMS architecture is presented. The MIPAMS section provides insights into the modules and services of the MIPAMS architecture that provide functionalities for the governance and protection of multimedia content. In order to illustrate the operation of this architecture, we present the results of some research and development projects in which the content management scenarios, previously described, have been implemented. Finally, a scenario for electronic publishing using mobile devices is presented. In this scenario the MIPAMS architecture is used for the registration and governance of digital information produced with smartphones.

Content Management Scenarios

Most of the literature refers to DRM as a means to restrict what users can do with content but in fact, DRM can be used in other contexts. For example, our research focuses on the "management" part; i.e., protection is not always necessary (for example in a trusted, or partly trusted, environment), but mechanisms to manage intellectual property are needed. We have identified several scenarios where DRM architectures enable the development of various kinds of applications on top of them, as detailed next.

DRM-enabled content access control. This scenario covers content registration, protection, search, licensing, authorization-based content access control, content storage and reporting. In this case, there is a need for an interface so that content creators can register and publish their content and determine and modify their offers. This functionality is provided by means of specific edition user

[1] This work has been partially supported by the Spanish Government through the project MCM-LC (TEC 2008-06692-C02-01).

applications or otherwise integrated in a web portal. Once content is registered, it can be linked from external sites so as to be able to license it through the mentioned portal, which means that the content promoted in external sites can include specific links towards the licensing portal. Moreover, apart from being linked from other sites, the portal itself would also be useful for promotion. In this business scenario, content is accessed by using DRM-adapted tools such as players and other rendering applications.

This scenario is illustrated in the following example. A news agency wants a solution for publishing, trading and distributing protected news. Content trade needs to support different licensing options, such as prices, time frames, territory, etc. Content access needs to be protected, controlled and reported. The news agency can make use of specific external services and a customized publishing and trading portal. Content access will be done through a DRM-enabled application.

Content Licensing. This scenario involves content registration, search, licensing and reporting. It is applicable to those cases where there are well established sites that deal with the promotion and collection of content, but for which licensing is not a part of their business model (e.g. Flickr, Picasa, Panoramio, YouTube, etc.). Although content can be directly accessed from those sites, it may be distributed under some restrictions that do not enable users to use it for free. This is the case when content is distributed, e.g., under copyright ("all rights reserved") or Creative Commons Non-Commercial models. In this scenario, there is a need for a trading portal, devised for formalizing the rights acquisition for personal or professional use. Content owners or rights holders are responsible for registering content in the trading portal and providing the link towards it. Content can be linked from external sites.

The following example describes the above scenario. A web portal dealing with specialized content (e.g., valuable medical images) wants to offer users the possibility to trade their content. The web portal defines some license templates that users can select when uploading their images. Content is automatically registered through external services and a link is provided from each image towards the trading portal for those users interested in licensing them for, e.g., publishing. Access to the images is managed by the web portal.

Content licensing and authorization-based content access control. This scenario involves content registration, search, licensing, authorization-based content access control, content storage (optional) and reporting.

When dealing with content storage, this scenario is useful for applications where users need to handle or modify content without restriction or when users do not want to be limited to using some specific DRM-enabled application. Although access to content is authorization-based, content is given unprotected to the purchasing users so that they can enjoy it without further DRM restrictions.

The next example describes the above scenario. A content distributor wants a solution for trading and distributing unprotected audiovisual content. Content trade needs to support different licensing options, such as prices, time frames, territory, etc. Content needs to be delivered unprotected, since it is to be transformed by its recipient to adapt it to different online and offline publishing formats. However, the content distributor wants to be sure that only those clients who own a license can download content. That is, content access needs to be controlled and reported. The content distributor can make use of specific external services and a trading portal. Content licensing and access can be done directly from the portal, after checking user licenses.

When content storage is not used, this scenario is devised for content providers or distributors who want to use their specific protection mechanisms and content management systems so that content is never stored outside their well-established systems. In such a scenario, when registering content, specific proprietary identifiers are used for identifying external content. Once objects are registered, rights offers can be published and licenses issued without restriction. Regarding the applications that access content, such as players and editors, content providers or distributors will have to design their own applications to manage the access to encryption keys and content from their systems or otherwise provide an API so that their content can be accessed from third-party applications.

Here is an example that illustrates the above scenario. A TV broadcaster wants to license his/her own productions. The TV broadcaster may make use of specific external services to register the content and define different licensing options (offers). However, content will never be stored outside his system. The TV broadcaster may develop his/her own trading portal that interacts with specific

services to consult the available offers and formalize the acquisition of an offer by issuing a license. Specific services can be used to check whether a customer has an appropriate license and enable content download (authorization-based content access control).

DRM-enabled content access for mobile devices. This scenario involves content registration, protection, search, licensing, authorization-based content access control, content storage and reporting. It is devised for limited capability devices. In some cases, the encryption strength being used should be limited so as not to be detrimental to the device's performance. In such cases, if content is already registered and protected, content has to be re-encrypted to deal with device limitations.

The following example illustrates the above scenario. A content distributor who is already using specific external services for trading and distributing protected content wants that content to be available for mobile devices. Since the device's decryption capabilities are limited, content may not use the same encryption strength as for PC-devised content or may be adapted to fulfill the device's requirements.

Intellectual property registry with external licensing. This scenario involves content registration, search and reporting. It is based on the use of registration functionalities, leaving content licensing to be tackled by professional external sites or services. In this scenario, there is only need for an intellectual property registry, proving content ownership and offering the possibility to link content to external sites that deal with its commercialization, as in, e.g., YouLicense, Getty Images, etc.

The next example describes this scenario. A web portal dealing with specialized content (e.g., valuable breaking news images) wants to offer users a powerful means for proving authorship. The web portal integrates external services into its registration process and automatically registers user content. A digitally signed document is available in the portal to certify authorship. The web portal may reach further agreements with specialized portals for offering licensing solutions to users.

The MIPAMS Architecture

MIPAMS (Multimedia Information Protection And Management System) is a service-oriented content management platform, developed by the DMAG (Distributed Multimedia Applications Group) (DMAG, 2011). It is mainly intended for applications where management of rights is needed.

The MIPAMS architecture is based on the flexible web services approach, as it consists of several modules and services which provide a subset of the whole system functionality needed for governing and protecting multimedia content. One of the advantages of having service-oriented DRM functionality is the possibility of decoupling it into different subsystems depending on the needs of the application that is going to be implemented, while being able to share the same common services between different applications with different requirements, thus reducing costs. MIPAMS encompasses an important part of the content value chain, from content creation and distribution to its consumption by end users.

Figure 1 depicts the MIPAMS architecture, for which we now provide a general overview of its components and the different services being offered.

The Content Service (CS) enables applications to upload and download digital resources such as audio or video files, text documents, etc. Those resources can be optionally encrypted on request, according to the available encryption mechanisms it provides. If encryption is selected, the protection keys will first be requested from the Protection Service and then registered through the same service, once encryption is performed. Content upload requires content to be uniquely identified. Since MIPAMS deals with single resource objects, the identifier associated to content will be the same one used for the object that contains it, and must be used as input argument. This identifier can be requested from the Object Registration Service prior to the content upload, or obtained from an external application using MIPAMS (depending on the scenario).

The Object Registration Service (ORS) enables applications to request a digital representation of content and metadata (i.e., digital objects) to be generated and registered in the system. Content and metadata are packaged together following the MPEG-21 Digital Item (ISO/IEC, 2005) approach. Once registered, objects are digitally signed by the ORS so that they can be checked for authenticity and

integrity. The ORS also provides unique identifiers for those applications that need to upload content to the CS, as already explained.

The License Service (LS) deals with rights offers and the issuance of licenses. Rights offers are set up by content creators or rights holders after registering content. They include the rights being offered for acquisition by other users and the conditions applicable to those rights. License issuance refers to the process by which a license is generated as the result of a rights purchase, acquisition or the direct grant by a rights holder to a user of a set of rights. Licenses are expressed using MPEG-21 Rights Expression Language (ISO/IEC, 2004).

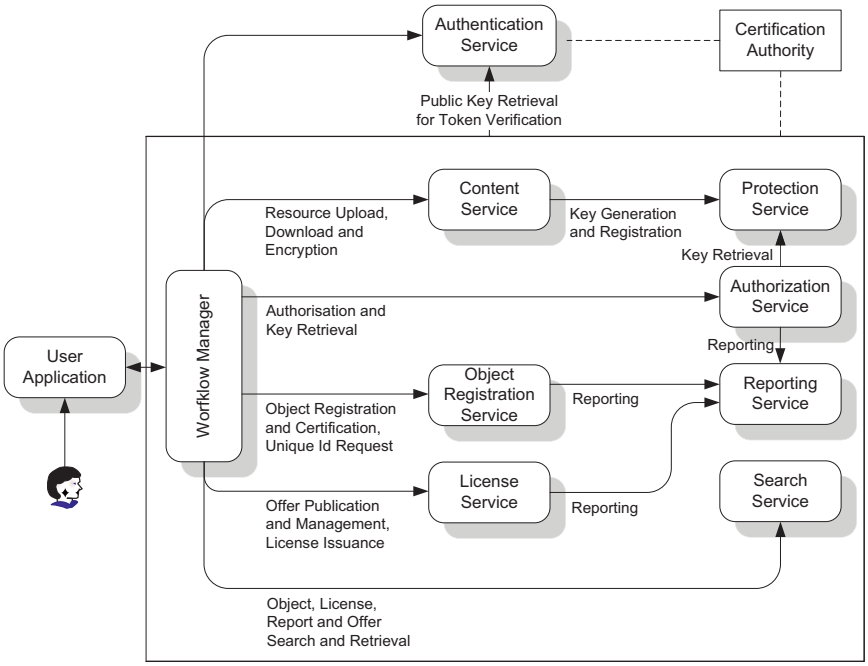

Figure 1. MIPAMS architecture

The Authorization Service (AS) checks whether a user owns any appropriate license that grants him the right to perform a requested action (e.g., play) on a digital object. The authorization is based on the mechanism defined in (ISO/IEC, 2004). The AS shares the access to the license repository with the LS. If the user is able to perform the action and the requested content is encrypted, the AS will retrieve the encryption keys from the Protection Service and return them to the requesting application. This is the only means for accessing encryption keys, which is performed as an atomic operation.

The Protection Service (PS), as introduced before, generates encryption keys upon request, registers encryption keys associated to uniquely identified content and provides the encryption keys for protected content to the AS. When using MPEG-21 Intellectual Property Management and Protection (ISO/IEC, 2006-1) scheme and descriptors, the PS also offers the possibility to download the protection tools being used by those applications that might be out-of-date.

The User Application (UA) is the player, edition tool, browser or any other means managed by the user to deal with the DRM functionality, such as registering and accessing protected contents. The UA may have an internal trusted module or intermediary to enforce DRM, which could consist of a secure local repository for licenses, protection information, offline operation reports and other critical data. It may be responsible for estimating tool fingerprints, require offline authorizations, unprotect content, track offline operations and manage content protection information.

The Workflow Manager (WM) may be an integral part of the UA or otherwise be located in the server part (e.g., web portal, brokerage service) to reduce the UA complexity. It can be seen as a broker whom the UA requests to perform different operations, such as object registration, content upload, rights offer management, license acquisition, authorization, etc.

The Search Service (SS) enables applications to perform accurate searches amongst metadata in the MIPAMS system. That is, it is the front-end for requesting any information present in MIPAMS

services databases. Thus, it can be used for searching content, licenses, offers or reports or a combination of them.

The Reporting Service (RS) collects usage reports regarding the registration of objects, the issuance of licenses and the authorizations being performed. It is also capable of building standards-based representations of those reports, such as MPEG-21 Event Reports (ISO/IEC, 2006-2). Those reports may be used for computing statistics as well as for billing or tracking purposes.

The Authentication Service (ATS) is needed to authenticate the identity of users. It generates SAML (Security Assertion Markup Language)-based tokens (OASIS, 2005) that identify MIPAMS users. Any service in the MIPAMS architecture will require a token argument to be provided in order to authenticate users. Tokens are digitally signed by the ATS, so that they can be checked for authenticity and integrity by the receiving service. Moreover, the ATS deals with user registration and management (i.e., personal data modification, user account deactivation, etc.).

Finally, there is a need for having a recognized Certification Authority (CA), which issues credentials for the different Components and Actors in the system, such as X.509 certificates and private keys for the different architectural components.

Implementation of Content Management Scenarios with MIPAMS

In this section we present the results of some research and development projects where we have implemented the usage scenarios previously identified as using MIPAMS services and modules.

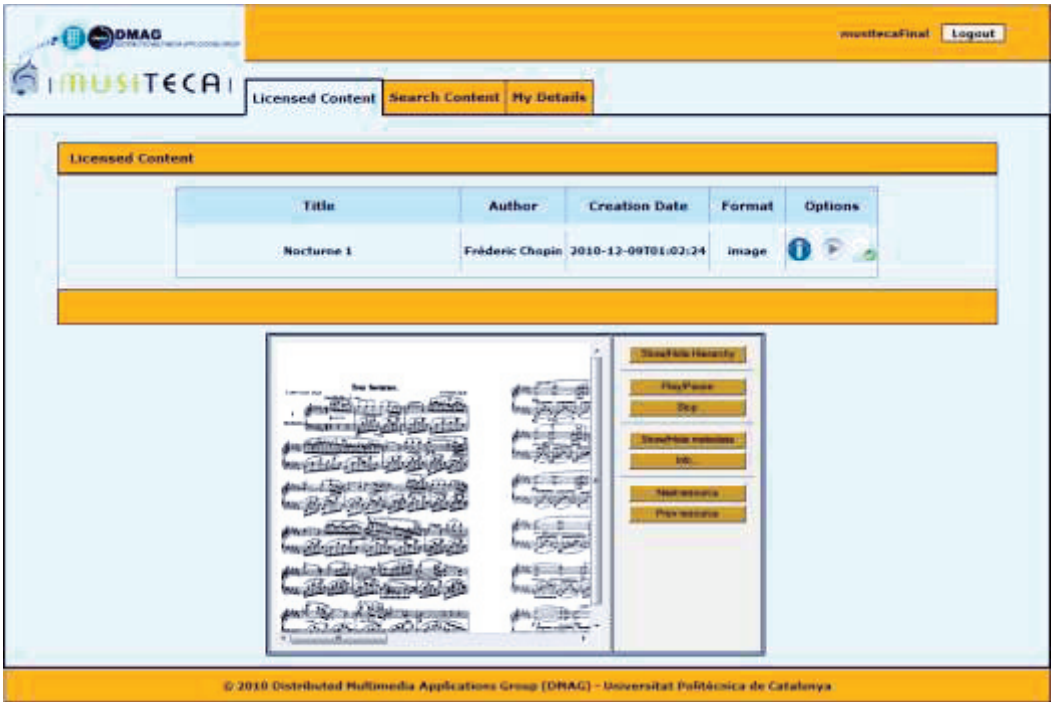

Figure 2. Protected rendering in a specific DRM portal in Musiteca

DRM-enabled content access control

This scenario has been implemented in Musiteca (Musiteca, 2008), a research project funded by the Spanish Administration. In this project, we have used some of the services making up MIPAMS (LS, RS, ATS, CS, ORS, SS and CA) to implement an electronic content publishing and trading platform. Access to the Musiteca repository is through a web portal that enables the registration of content, the definition of different licensing options (offers), the purchase of content (licensing) and access to the content purchased after checking whether the user is authorized. Content access is through a DRM-enabled application, while any action in the system is registered through the RS. Figure 2 shows a screenshot of the portal, where content is being rendered.

Content Licensing

This scenario has been also implemented in Musiteca. Figure 3 shows how content is linked from an external site, the Musiteca base on Freebase, which holds information about musical content in the Musiteca project, to a specific trading portal where content can be licensed and accessed.

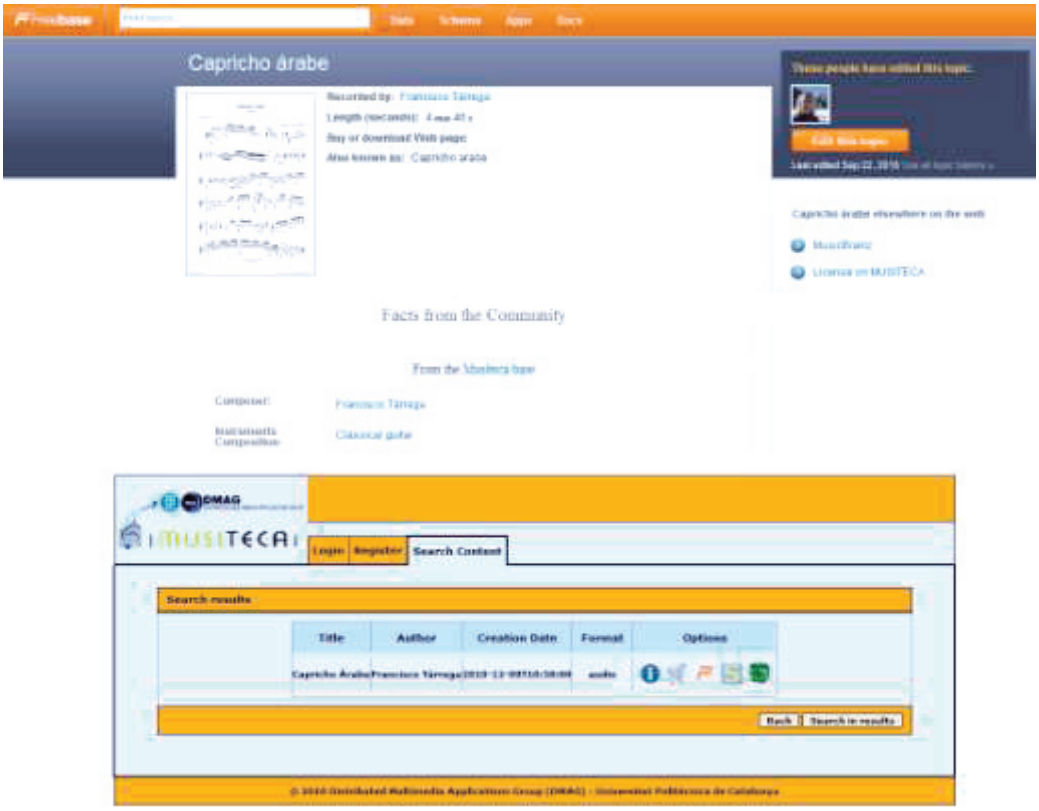

Figure 3. Licensing link from Freebase to a specific trading portal in Musiteca

Content licensing and authorization-based content access control

This scenario has been implemented in the Intellectual Property Operations System (IPOS) (IPOS, 2011), a content lifecycle management system resulting from several software developments carried out by the DMAG under different contracts with NetPortedItems (NetPortedItems, 2010), a Spanish SME company. IPOS provides content authors with the possibility of registering, publishing their work and defining how other users can license the work for deriving new content. This information is described using LS licenses, where we have added a special condition called Rights over Derivatives (ROD) (Torres, 2009). This condition indicates the percentage of the income that someone gets from a derivative work which will be owed to the original author. When an author creates derived content from an existing work and gets any revenue from it, IPOS follows the chain of work back, calculates the share for each author from the ROD condition in the licenses and creates a report for each author informing him of this fact. Reports can be consulted at established time periods to give each author the corresponding revenues. This system makes use of all MIPAMS services through a dedicated portal. Figure 4 shows a sample screenshot of the authorization-based rendering application.

This scenario has also been implemented in CulturaLive (CulturaLive, 2009), a research project funded by the Catalan Administration. In this project we have integrated, using Web Services, MIPAMS LS, AS and RS into an existing system offered by another project partner (VSN, 2011) that provides audiovisual content to be broadcast live through Digital Terrestrial Television (DTT) by televisions participating in the project. With our modules, content purchases can be tracked since we register each license acquisition and authorization result (positive or negative) into a reporting database. This database can later be consulted for billing purposes. It is worth noting that digital content to be broadcast is not managed by MIPAMS but directly by the different TV channels and

SMEs in the project consortium. This gives an idea of the integration capabilities of the MIPAMS platform.

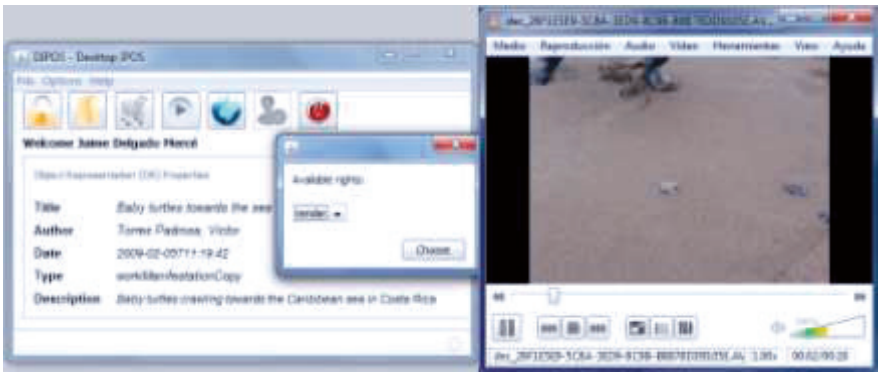

Figure 4. Content access and unprotected rendering in IPOS

DRM-enabled content access for mobile devices

This scenario has been implemented in some projects of our research group (e.g. AXMEDIS (AXMEDIS, 2004), but also in other projects) in a slightly different way. In such projects, the modules involved in the authorization of user actions were located inside the mobile device. In this way, when the user wanted to consume some content, the license for authorizing this action was inside the mobile. This was done to avoid calling external services, as it involved a phone call or data transaction that might involve a non-negligible cost for the user. Moreover, mobile devices used a specific licensing schema (OMA DRM, 2010) addressed to devices with limited processing and communication capabilities. Currently, since smartphones and high capacity mobile devices are gaining relevance and current telecommunications companies are adopting competitive pricing policies for mobile users (e.g., flat data fees), the solutions being implemented might be reconsidered.

To implement this scenario with MIPAMS, if content is already registered and protected using a protection mechanism not compatible with the device, the intermediary would be responsible for decrypting content and re-encrypting it to deal with the device limitations. Otherwise, if content is only to be used by limited capability devices, it should be encrypted using the suitable protection mechanism when uploaded to the CS.

Intellectual property registry with external licensing

Figure 3 (the lower part) shows how content could be linked from the MIPAMS-based intellectual property electronic registry developed in the Musiteca project towards external specialized licensing portals. Some examples (not used in this project) are YouLicense (YouLicense, 2011) or Getty Images (Getty Images, 2011). Content would be registered and accessible to search, while the shopping cart icon would redirect the user to a specialized and external licensing service.

Scenarios where Mobile Devices are needed

Among the different scenarios presented in the previous section, there is one specifically for limited capability devices. Inside this category of devices we include smartphones and other similar devices with internet connection that usually have an integrated camera. The difference between those devices and personal computers is mainly their dimensions (smaller), their processing capabilities (lower) and, maybe the most important, that they do not have a qwerty keyboard or mouse to interact with the user, but a touch screen or a keypad.

In this case, if we want to register new content for publishing it electronically, using one of these devices could be a laborious and slow process. Nevertheless, it could be interesting in some specific cases that we are going to describe in this section. We will make use of the MIPAMS architecture to offer these services for smart devices.

First of all, the user application has to be prepared for use in those devices. For the case where the application is a web portal, there is not much problem, as current devices have integrated browsers

which are able to manage web forms and show web pages. The only restriction in some cases is the screen size, but this can be avoided with a suitable design.

There is a scenario where mobile devices can be useful for electronic publishing, although the quality of the recorded images or videos is obviously lower than those produced with professional devices. In this scenario, what is most important is the "opportunity" of the content in the image or the video, that is, to be at the right place at the right moment, independently of the content quality. This is especially true for sudden events, for example a building collapse, a natural disaster or some celebrity being caught in an embarrassing or funny situation. In such cases, the images may even have economic relevance, as they could be published in online newspapers or in the gossip news. Therefore, the author may register the content for different purposes: to try to get some revenue or just for later attribution.

MIPAMS modules involved in this scenario and their interaction are depicted in Figures 5 and 6. Figure 5 shows the registration of the content, image or video in the registration portal. Figure 6 shows how the author can create some offers of the content to get revenue. If the author only wants to be able to prove authorship, registering the content would be sufficient. In this case a Creative Commons attribution license would be suitable as long as the author waives his right to any potential future revenues (Creative Commons, 2011).

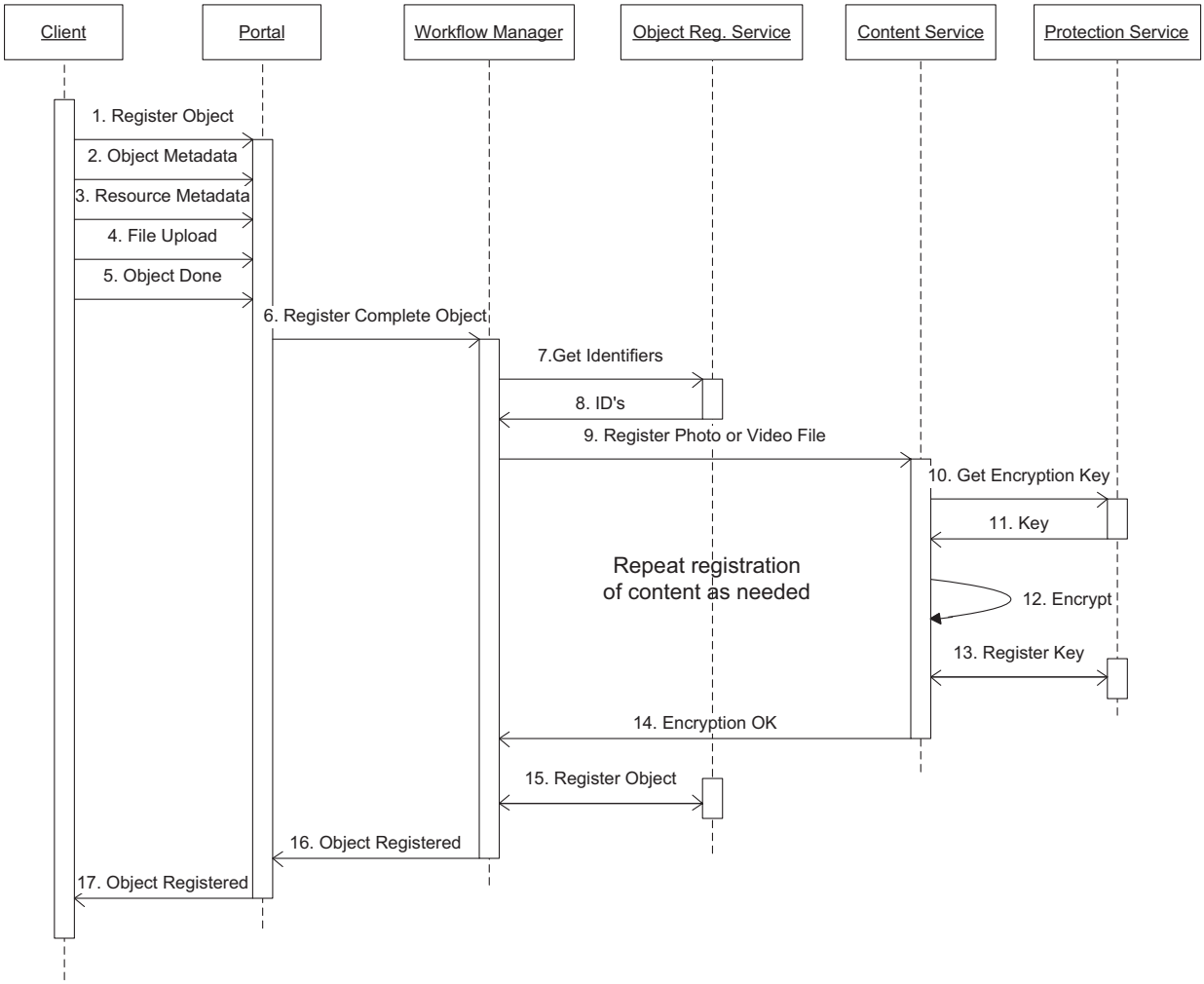

Figure 5. Registration of photo or video

The steps involved in content registration shown in Figure 5 are the following:

1. User starts registration of content in the portal by means of the client application. The client application can be a browser or a specific application for mobile devices, where some information could be predefined to facilitate the registration.
2. User fills out a form with all metadata associated to the complete digital object.
3. User fills out several forms (one for each image or video, that is, each resource) with metadata associated with each resource.
4. User uploads the file containing each resource.
5. User indicates that all object information has been inserted and the registration process needs to continue.
6. The portal sends all information to the Workflow Manager (WM) module, which will call the corresponding service for storing the object.
7. WM requests identifiers from the Object Registration Service (ORS).
8. ORS sends the identifiers requested to the WM, one for the object and one for each resource file (even if they have not yet been uploaded).
9. WM sends resource to the Content Service (CS)
10. If user has requested encryption of the resource, CS asks for encryption keys from the Protection Service (PS).
11. PS returns the keys for encryption algorithm and key length specified by CS.
12. CS encrypts and stores the file with the given key.
13. CS registers the encryption key in the PS for permitting later decryption.
14. CS sends WM notification of correct content storage and encryption. Steps 9 to 14 are repeated for each resource uploaded by the user. If no resources are uploaded, these steps can be done later.
15. When all resources are properly uploaded and encrypted, WM requests from ORS registration of the complete object, which is digitally signed to guarantee digital object integrity. The format used for storing the object is the MPEG-21 Digital Item.
16. WM sends notification of object registration to the portal.
17. Portal informs the user that the object has been properly registered.

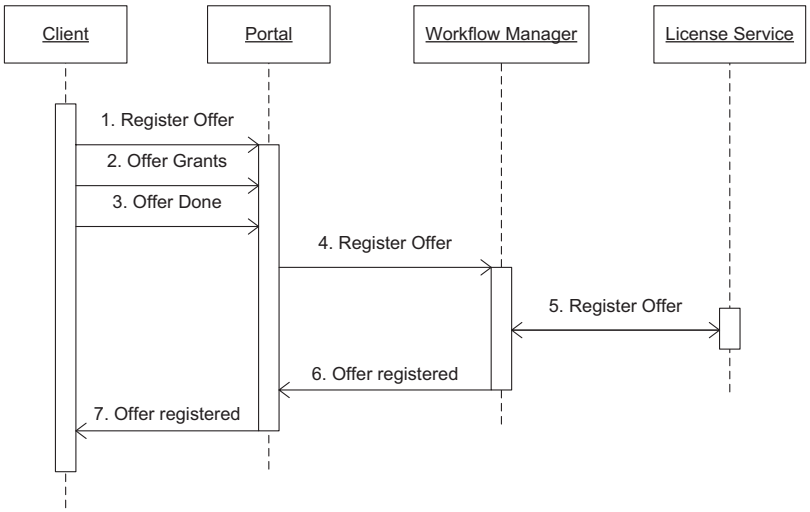

Figure 6. Creating some offers to get revenues

After object has been registered, the user can create some offers to profit from the images or videos she has registered. The steps involved in offers creation shown in Figure 6 are the following:

1. User asks Portal for offer registration.
2. User has to insert the different sales conditions offered for the registered object. These conditions include what can be done with the content (play, print, etc.) together with conditions such as territory, number of times one can perform the action or terms of payment.
3. User indicates to the Portal that the offer is complete.
4. Portal sends the offer to the WM.
5. WM sends offer information to the License Service (LS) which checks whether everything is correct and stores the offer.
6. WM informs the portal of offer registration.
7. Portal informs the user of offer registration. From that moment on, the registered content can be sold with the purchase conditions indicated by the user. It is possible to create an offer saying that the digital content is free, just to guarantee author's attribution.

Once registered, other users are able to purchase the object. At that moment, a license is created based on the selected offer. This part of the scenario corresponds to the *Content licensing and authorization-based content access control with content management and without protection* business model already described in the scenarios section, since content needs to be further processed prior to being published.

Conclusions

Mobile devices may become crucial when they are the only available means for recording information (e.g., taking a photo or recording a video) of an unexpected event of any type.

In this context, we have presented the MIPAMS architecture, developed by DMAG, devised for the management and secure distribution of multimedia content. The operation of MIPAMS architecture has been presented by means of the results of research and development projects in which different content management scenarios have been implemented.

We have illustrated different scenarios where MIPAMS has proved to be useful for copyright preservation and electronic publishing, some of them beyond traditional DRM applications.

Finally, we have focused on limited capability devices, including a detailed description of how MIPAMS registration, licensing, trading and authorization-based access control functionalities may help to protect the author's copyright, ensuring attribution and easing the exploitation of his economic rights. A detailed analysis of the interaction between different components of the architecture has been also provided.

References

AXMEDIS (IST-2004-511299). (2004-2008). Automating Production of Cross Media Content for Multichannel Distribution. Retrieved May 2, 2011 from http://www.axmedis.org.

Creative Commons licenses. (2011). Retrieved May 2, 2011 from http://creativecommons.org/licenses/.

CulturaLive Research Project (2009REGIÓ 00024). (2009). Generalitat de Catalunya.

Delgado, J., Torres, V., Llorente, S. & Rodríguez, E. (2011). Rights management in architectures for distributed multimedia content applications. *Trustworthy Internet*. Heidelberg: Springer. Publication pending.

Distributed Multimedia Applications Group (DMAG). (2011). Retrieved May 2, 2011 from http://dmag.ac.upc.edu

Getty Images. (2011). Retrieved May 2, 2011from http://www.gettyimages.com/.

Intellectual Property Operations System (IPOS). (2011). Retrieved May 2, 2011 from http://dmag1.ac.upc.edu/IPOS

ISO/IEC. (2005). ISO/IEC IS 21000:2 – Part 2: Digital Item Declaration.

ISO/IEC. (2004). ISO/IEC IS 21000:5 – Part 5: Rights Expression Language.

ISO/IEC. (2006-1). ISO/IEC IS 21000:4 – Part 4: Intellectual Property Management and Protection Components.

ISO/IEC. (2006-2). ISO/IEC IS 21000:15 – Part 15: Event Reporting.

Musiteca Research Project (TSI-020501-2008-117). (2008). Ministerio de Industria, Turismo y Comercio (Subprograma Avanza I+D). Retrieved May 2, 2011 from http://musiteca.freebase.com/.

NetPortedItems S.L. (2010). Retrieved May 2, 2011 from http://www.digitalmediavalues.com/.

OASIS. (2005). Security Assertion Markup Language (SAML). Retrieved May 2, 2011 from http://saml.xml.org/

OMA DRM: Open Mobile Alliance Digital Rights Management. (2010). Retrieved May 2, 2011 from http://www.openmobilealliance.org/technical/release_program/drm_v2_1.aspx.

Torres, V., Delgado, J., Maroñas, X., Llorente, S., & Gauvin, M. (2009). A web-based rights management system for developing trusted value networks. In *Proceedings of the 18th International World Wide Web Conference Developer's Track, April 20-24, 2009, Madrid, Spain* (pp. 57-59). New York: ACM. Retrieved May 2, 2011 http://upcommons.upc.edu/e-prints/bitstream/2117/7776/1/www09dev_proceedings.pdf.

Video Stream Networks (VSN). (2011). Retrieved May 2, 2011 from http://www.vsn-tv.com/es.

YouLicense. (2011). Retrieved May 2, 2011 from http://www.youlicense.com/.

Next-step Digital Publishing Tools and Practices

Federico Ruberti
FakePress, Via A. Toscani 27, 00152 Rome, Italy. Email: federico@fakepress.net

Luca Simeone
FakePress, Via Cutilia 17, 00183 Rome, Italy. Email: luca@fakepress.net

Abstract: *The rapid growth and proliferation of digital devices combined with cultural and technological changes suggests that in the near future we will see a deeply changed scenario in the publishing industry. Mobile technologies in the hands of a networked society should cause those players who wish to maintain a crucial role in the content industry, such as publishers and editors, to reconsider the way in which content is created, distributed, shared and consumed nowadays.*
Keywords*: Digital publishing; cross-media design; emergent narratives; location based media.*

> ***"Publishers have to become more reader oriented in a marketing and trend finding/ setting way rather than in a direct to consumer selling way".***
>
> Markus Dohle, CEO of Random House Interviewed by Mike Shatzkin, September 2010 (http://www.publishingtrends.com).

Introduction: A Confusing Scenario

Nowadays the publishing industry seems to face a period of uncertainty, due to an unpredictable evolution of its business model and the vagueness of its role in a deeply changed society. While in the past the work of the publisher could be synthetically described as reviewing manuscripts, editing, marketing and distributing physical books through booksellers around the world, as the business model began to shift, hard-pressed by cultural changes and emerging technologies, many players began to wonder which of the publisher's roles could still be relevant in the near future. Chain narratives, peer reviewing and collaborative content creation, far from the earlier experiments, have reached a considerable level of maturity and reader appreciation; "Self publishing" and "print-on-demand" have become a viable option for a growing niche market; the continuing march toward digital publishing models is making significant changes in business models and processes faster than the publishing industry has ever seen before (Book Industry Study Group, 2009).

According to what Jim Milliot reported in *Publishers Weekly* in 2009 (Milliot, 2009), in the USA, the number of new and revised titles produced by traditional production methods fell 3% in 2008, to 275,232, but the number of on-demand and short run titles soared 132%, to 285,394. Since 2002, production of on-demand titles has grown 774% compared to a 126% increase in traditional titles. Apart from some extreme technological experiments in book production, such as Espresso Book Machine (On Demand Books, 2011), a large photocopier machine, launched in London 2009, that prints and binds books on demand in five minutes, the rapid development of print technologies has greatly contributed to the spread of the self-publishing phenomenon, affecting both economics and processes in book production. Players such as Lulu.com have made it possible to self-publish a book at a cost that is insignificant if compared to the past. Furthermore, the rapid diffusion of e-readers and tablet PCs, combined with the success of e-books, has dramatically reduced costs, transforming self publishing from a cheaper way of publishing to a profitable form of publishing in all respects, including the royalties earned by authors for each copy sold.

The efficacy of most marketing and communication strategies put in place by editors and publishers for the promotion of books and authors is also questioned. As in the old days when there were few mechanisms for distribution, and fewer chances of bringing books to the attention of relevant book reviewers, in the age of the Internet readers have discovered that they don't necessarily care about what reviewers say about a book if it covers a topic in which they are interested. The economic impact

of these dynamics on the book industry should be another matter of concern for most editors and publishers. Since 2006 Vasant Dhar, professor of information systems at New York University Stern School of Business, has shown through a sound analysis (Dhar & Chang, 2008) of the economic impact of user-generated content for the music business, how the music industry has failed not because it was unprepared to recognise self organised phenomena within the consumer base but because instead of focusing on providing their customers value and reasonable rights of usage, the music industry became obsessed with preventing piracy, and it cost them dearly. In this sense publishers should adopt a more market-focused business model that welcomes technological innovation instead of continuing to be disconnected from their rising consumer base, and in so doing underestimating the power of the cultural shift that has been brought about by emerging technologies.

Readers are increasingly self organised as communities of interests that share information, opinions and comments, trusting each other more than advertising and reviewers' opinions. The importance acquired in recent years by User Generated Content has exceeded that of traditional marketing and communications strategies in the publishing industry as in consumer goods and retailing, where in an Internet search carried out on 20 of the world's most famous brands, 25% of the results were links to User Generated Content. On the web 34% of bloggers express comments and opinions on products and brands to a public of whom 78% trust their recommendations (Qualman, 2009). The groundswell, that is, the trend of obtaining information and goods from other individuals rather than from businesses and institutions, assumes, along with the evolution of mobile technology, a disruptive potential which publishers must face. "As powerful as it is, technology is just an enabler. It's the technology in the hands of almost always connected people that makes it so powerful" (Li & Bernoff, 2006).

In relation with the news industry, this phenomenon has recently been analysed by the Pew Internet and American Life Project (Purcell, Rainie, Mitchell, Rosenstiel & Olmstead, 2010) according to which people's relationship to news and information is becoming increasingly portable, personalized, and participatory: 33% of mobile phone owners access news on their phone, 28% of Internet users customise their home page with news from third party sources, 37% of Internet users contribute to the creation of news, comment about it, or disseminate it via blogs and social media web sites. Marketing and communication strategies that have characterised the positions taken by the publishing framework within the complex system of promotion of books and authors, and the even more complex system that generates profits through book selling, are rapidly making way for social media and user generated content, so that editors and reviewers may influence readers' preferences much less than in the past.

Collaborative content creation, as both technological solution and cultural attitude, is rapidly spreading not only in the academic sector. Looking at some recent experiments, such as Booki,[1] The Book Oveni[2] The Book of MPub[3] or CommentPress[4] it could be argued that there are numerous "communities of interest" that are increasingly involved with the technological solutions and environments that allow individuals to cooperate and share comments, access, reuse and remix content, and that the issue of formats and size (printed book, print, epub, pdf, HTML5, html and so on) through which the content is released, looks like a purely functional issue; the focus is on collaboration, and formats are relevant only as a way to make content accessible.

Giovanni Ragone (2005) talks about the fourth generation of publishing, where the web serves as an increasingly important route for communication as a source, as an archive and as an environment in

[1] At Booki, people are encouraged to write a book by themselves or with others, discuss views, seek assistance, translate or reuse content and publish completed works (http://www.booki.cc).

[2] The Book Oven, an open source book publishing platform, helps teams of people turn manuscripts into finished books, and then publish them generating clean, well-formatted books in multiple outputs: .epub, print-ready PDF, InDesign-ready XML, and so on. It is built for writers, editors, proofreaders, designers and small presses (http://bookoven.com/splash).

[3] The Book of MPub, launched April 16, 2010, curates research and critical thinking from students in the Master of Publishing program at Simon Fraser University. In doing so, it makes a contribution to a collective discourse on innovative technologies in publishing—epublishing, new business models, and crowd sourcing and social media. The Book of MPub furthers discussion in three formats: blog, ebook and the classic, ever-evocative print form (http://tkbr.ccsp.sfu.ca/bookofmpub).

[4] CommentPress is an open source theme and plug-in for the WordPress blogging engine first released by the Institute for the Future of the Book in 2007 aiming to allow readers to comment paragraph by paragraph in the margins of a text turning a document into a conversation. It can be applied to a fixed document (paper/essay/book, etc.) or to a running blog (http://www.futureofthebook.org/commentpress).

which readers organize their own newsrooms; the web 2.0 is like a non-place where contents and products are always available and accessible, exploiting the characteristics of a digital text (multimedia, hypertext, interactivity). The new web by its nature subverts the traditional roles of producer and consumer of content, creating an environment where people unknown to each other share information and contents on the basis of common interests.

Changes that are occurring suggest a "cognitive" revolution linked to the network that will have an impact on internal processes throughout the publishing industry.

Conclusion

In the next few years, digital devices, interactive technologies, collaborative platforms, advanced distribution systems, augmented reality, will revolutionise the publishing industry. The sector will see the creation and dissemination of publications, cross-medial, multi-authored, open-ended, and intimately associated with the web and community of "professional readers" using mobile technologies, location-based platforms, and interactive design. Publishing houses, then, will have to reconsider their role and consequently update their processes and products to keep up with social, cultural and technological changes that have profoundly modified the way in which digital content is produced, distributed, displayed and accessed.

In a changed scenario, publishers may still play a significant role as a bridge between authors, content and readers, sustaining communities and creating a solid technological framework that enhances multi-authorship, cross-medial distribution, interactivity and knowledge dissemination.

Aiming to reconsider the role of publishing houses and to identify possible ways in which they would continue to represent a fundamental cog in the publishing engine, it could be suggested that successful publishers may be those that:

- enhance interactivity, so that content can be updated, extracted, remixed, re-contextualized, distributed, reviewed by both readers and authors;
- design and develop a robust technological framework that inspires a complex range of user experiences and new forms of content dissemination, through the web and mobile platforms;
- create an integrated technological environment that transforms standard multimedia content (text, images, videos...) into its cross-medial version;
- actively contribute to building communities that involve authors and a group of readers interested in a topic, encouraging projects that aim to build and nurture these communities, experimenting and developing new ways of visualizing, consuming and sharing knowledge developed around a theme;
- sustain multi-authorship, leaving contents to be hosted in environments that welcome multiple voices, points of view and layers for interpretation, at any time and place; and
- produce publications that address multiple media in harmonious ways so that contents may be accessed on paper, web, mobile phones, architectures, objects, and bodies, by using ubiquitous and location-based technologies.

References

Book Industry Study Group. (2009). *Book industry trends 2009*. New York: The Book Indsutry Group.

Dhar, V. & Chang, E. (2008). Does chatter matter? The impact of user-generated content on music sales, *Journal of Interactive Marketing* (February 2008). CeDER Working Paper No. 07-06. Retrieved May 28, 2011, from SSRN: http://ssrn.com/abstract=1113536.

Li, C. & Bernoff, J. (2006). *Groundswell: Winning in a world transformed by social technologies*. Boston: Harvard Business Press.

Milliot, J. (2009). Number of on-demand titles topped traditional books in 2008. *Publishers Weekly*, May 19, 2009. Retrieved May 12, 2011 from http://www.publishersweekly.com/pw/by-topic/industry-news/bookselling/article/9038-number-of-on-demand-titles-topped-traditional-books-in-2008-.html

On Demand Books. (2011). The EBM Hardware. Retrieved May 12, 2011 from http://www.ondemandbooks.com/hardware.htm

Purcell, K., Rainie, L., Mitchell, A.; Rosenstiel, T., & Olmstead, K. (2010). *Understanding the participatory news consumer*. Washington, D.C.: The Pew Research Center for the People & the Press.

Qualman, E. (2009). *Socialnomics: How social media transforms the way we live and do business*. New Jersey: Hoboken.

Ragone, G. (2005). *L'editoria in Italia. Storia e scenari per il XXI secol*o, Napoli: Liguori.

Further readings

Hendrickson, C., Iaconesi, S., Persico, O., Ruberti, F., & Simeone L. (2010). *REFF. The reinvention of the real through critical practices of remix, mash-up, re-contextualization, reenactment*. Rome: Derive Approdi.

Stein, B. (2008). A unified field theory of publishing in the networked era. In *Proceeding of the 2008 ACM workshop on research advances in large digital book repositories. Napa Valley, California USA.*

Sophie 2.0 and HTML5: DIY Publishing to Mobile Devices

Kalin Georgiev

Astea Solutions AD, Sofia, Bulgaria. Email: kgeorgiev@asteasolutions.com

Nicholas Matelan

Project Manager, Sophie 2.0, Los Angeles, CA, United States of America.
Email: nmatelan@matelan.com

Ludmil Pandeff

Astea Solutions AD, New York City, NY, United States of America.
Email: lpandeff@asteasolutions.com

Holly Willis

Institute for Multimedia Literacy, University of Southern California, Los Angeles, CA, United States of America. Email: hwillis@cinema.usc.edu

Abstract: *For an e-book to reach its maximum audience, it needs to be accessible on all modern reading devices in a consistent way. This work compares existing technology choices for publishing e-books for iPad and Android tablets. The research has been conducted in the context of Sophie, software for reading, writing, annotating and sharing networked books, with attention to the ongoing evolution of the software necessary to keep it relevant in a swiftly changing technology landscape. In particular, the article charts the development of the ability to create Sophie projects specifically for a new generation of mobile devices and tablets in support of a generation of students, scholars and publishers who expect easy collaboration, participation and interaction, and tools that allow for media-rich, mobile experiences.*

Keywords: *Sophie; mobile; e-publishing; HTML5; iPad Readers; Android Readers; E-book formats.*

Introduction

Sophie (Georgiev & Sredkov, 2010) is software for reading, writing, annotating and sharing networked books. With considerable support from The Andrew W. Mellon Foundation, Sophie has become a unique tool, allowing authors with normal computer skills to create professional e-books quickly and easily, without sacrificing the customary habits and practices unique to traditional reading and writing.

The primary users of Sophie at this point in 2011 are undergraduate students and faculty members on college and university campuses, where the drive to update learning by making it more collaborative and participatory coincides with the desire for new kinds of scholarly communication attuned to the networked interactions that have transformed contemporary culture. As many have pointed out (e.g., Davidson & Goldberg, 2009), it is imperative that learning institutions understand a generation of students accustomed to participatory learning and networked interactions. Similarly, it is incumbent upon software developers, designers and those who expect to use new tools to work together to create applications suited to collaboration, interaction, and transformation. This is not a trivial undertaking, nor is it one that should be relegated simply to software developers. As Drucker (2009) points out, "The design of digital tools for scholarship is an intellectual responsibility, not a technical task". This paper illustrates how Sophie must now continue to evolve, working toward fulfilling its potential to allow users to create e-books that they can deliver to new reading devices, and explains the work undertaken to achieve that goal.

Continuous Evolution

Since its earliest incarnation as software released in the “Squeak” Smalltalk platform in 2008, Sophie has morphed in response to its sense of "intellectual responsibility" to a changing culture. When it became clear that Sophie required an extended user base in the open source community if it was to remain viable, the software was rewritten in Java to connect with that community. When it became clear that Sophie books composed as stand-alone objects no longer suited the desire for networked reading, and when notions of collaborative authoring moved from fantasy to reality, Sophie kept pace, and development focused on the Sophie Server, designed to allow users to collaborate easily and efficiently in real time.

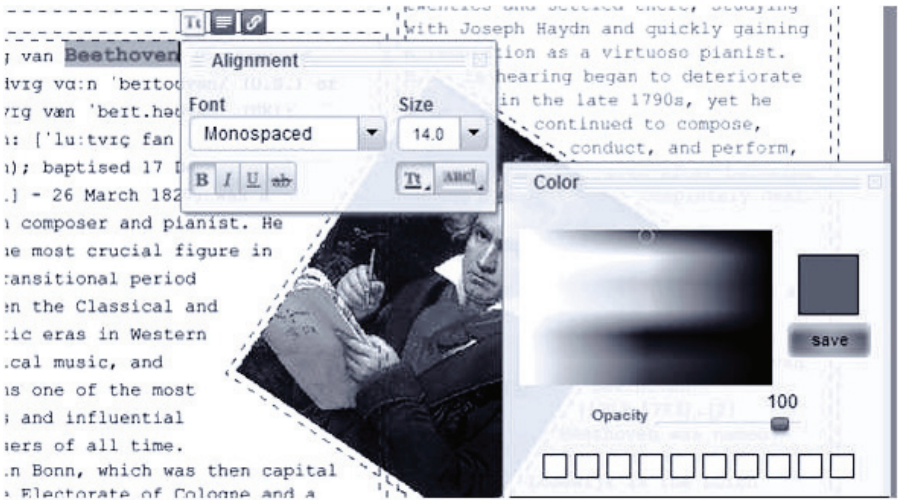

Figure 1. A glimpse of Sophie’s interface

Now, Sophie faces a similar juncture. Sophie must adapt to the recent explosion of interest in mobile readers such as the iPad and the Nook, and it must support publishing on a broad scale, offering the publishing community a robust and innovative tool for disseminating media-rich, networked reading experiences on a wide range of devices. To this end, the Mellon Foundation funded an initiative titled “Sophie 2.0: From Projects to Publishing,” which includes support for the creation of six exemplary book projects at universities across the USA, along with funding for two marketing projects dedicated to investigating Sophie's potential role in publishing. The first marketing project was a brief survey designed to gauge the impact on publishers of adding a feature to Sophie for exporting books in a format suitable for reading on an Apple iPad; the second was a longer survey (due for completion in second quarter 2011) to identify further functionalities desired by publishers.

While the iPad Survey was being planned, however, the industry continued its quick evolution, easily evident at the 2010 Consumer Electronics Show in Las Vegas, where tablets and networked reading devices were among the most popular – and prevalent – technologies presented. Low-power and low-cost processors that are robust enough to support networked activities have contributed to the proliferation of tablets and e-book readers, and dozens of manufacturers internationally have produced devices. Further, the Google Android operating system-based readers gained significance in the fourth quarter of 2010, with Gartner, a leading technology research company, stating in September 2010 that “the worldwide mobile operating system (OS) market will be dominated by Symbian and Android, as the two OSs will account for 59.8 percent of mobile OS sales by 2014” (Gartner, 2010). Indeed, this shift in turn prompted the Sophie team to revise the “iPad Survey” to become the “iPad and Android Survey.” Finally, yet another example of the rising significance of a new range of mobile devices was recently reinforced by the *2011 Horizon Report* (Johnson, Smith, Levine & Haywood, 2011), published jointly by the New Media Consortium and EDUCAUSE, which lists electronic books and mobile devices as the two most significant technologies destined to impact higher education in the coming twelve months.

The iPad and Android Survey

The perceived need for iPad and Android as primary publishing options was verified in several ways, the primary one being a survey of publisher interest in the platforms. The brief survey, conducted in Fall 2010, was focused on a single purpose: providing data for making an informed decision whether or not to extend Sophie functionality based on the importance of the iPad and Android to decision makers in the electronic publishing industry.

Publisher Corroboration by Survey Results

The market study was completed by the Morey Evans Agency to determine whether or not the Apple iPad reader and Google Android reader products are important to qualified experts in the book publishing industry. These experts were associated with organizations that included large publishers, small publishers, self publishers and academic presses. The study was a part of a Mellon Foundation funded project that includes the Exemplary Books project.

The Purpose of the Survey

The intent of the iPad-Android survey was very focused. It had the single purpose of providing data for making an informed decision on whether or not to extend Sophie functionality based on the importance of the iPad and Android to decision makers in the electronic publishing industry. Since it appeared that both iPad and Android readers could be served by the same mechanism, Android was added to test the validity of claims that Android was rising rapidly as an iPad competitor. Near term results for the survey were sought so that, if the results were positive, export extension could be added to Sophie to keep it in contention with other solutions in a rapidly evolving market environment. It is thought that having iPad and Android compatibility actually in work will give potential publishing adopters some assurance that these platforms will be supported and make the effort to attract new users far more productive. The need for an early determination drove many of the design decisions for the survey.

The Design of the Survey

The survey design was based on examining the importance of the iPad and Android as electronic book reading platforms to decision makers in publishing organizations. Targeting a specific population minimized coverage bias stemming from samples not representative of the population. Sources of coverage bias are typically a less carefully targeted respondent population, unreliable contact information, or lack of incentive or motivation to respond. It became clear that determining the correct non-biased discriminator for isolating decision makers was the overriding issue in crafting a meaningful survey. Thomas Minkus agreed to triage the collected job titles at publishing houses in order to give the best access to decision makers based on his twenty years of experience. Mr. Minkus is the Vice President for Emerging Media of the Frankfurt Book Fair, the largest book fair in the world, the major concentration of publishing executives each year. Using this guidance, it was determined that direct email addresses for specific people with these titles were difficult to obtain and that general email for a survey is typically discarded or not answered by the intended level of person. This tended to limit the initial sample in number, which in turn prompted a survey design aimed at improving response rate. Although the sample size was small, the sample represented the thought leaders in electronic publishing, and their opinions have greater weight than a larger, more general survey. The point was to target people who will actually make decisions about which books and tools to use in order to make a go/no-go decision for implementation of new export features for Sophie.

The practices guiding survey design were:

- Find contacts that are likely to have an interest;
- Keep the survey short and to the point;
- Make answering simple;
- Clearly state the purpose;
- Avoid coverage bias;
- Provide a respondent-friendly survey questionnaire;
- The question asked must be clear, non offensive and easy to respond to for the subjects under study.

These guidelines were employed with the counsel of USC professor and marketing expert David Weitzner.

The two questions asked were chosen to get a snapshot of the value of iPad and Android to this group in order to proceed with iPad additions to Sophie as quickly as possible, if warranted by results. The use of a simple, straightforward two-question survey without additional complication was employed to elicit the most reliable near-term measure. A simple one-page Website was used to gather results. This was done according to the proposal under the current Mellon project that specified this survey to support the need for the iPad for Sophie. (The second survey under the current grant is intended to gauge the need for a longer list of publishing capabilities and will be sent to a much larger group.)

Survey Raw Results

With a response of 48% from a small sample population (115) of publishing decision makers, the results of the survey are:

Question 1: Is iPad compatibility of your electronic books an important part of your publishing strategy?

Responses allowed were: Yes, No, I don't have an opinion.

Result was Yes 80.8%, No 19.2%; No opinion 0%

Question 2: Is Android compatibility of your electronic books an important part of your publishing strategy?

Responses allowed were: Yes, No, I don't have an opinion.

Result was Yes 73.1%, No 26.9%; No opinion 0%

Responders were all qualified professionals at all sizes of publishing organizations, with such titles as: Vice President International Sales, Publisher, Editor in Chief, Vice President Sales, and Publishing Director. Publishing organizations included: QA International, Hachette, Simon & Schuster, Bloomsbury, Penguin Group, Open Letter Books, Yale University Press, Equinox Publishing, Wiley, and Galileo Books.

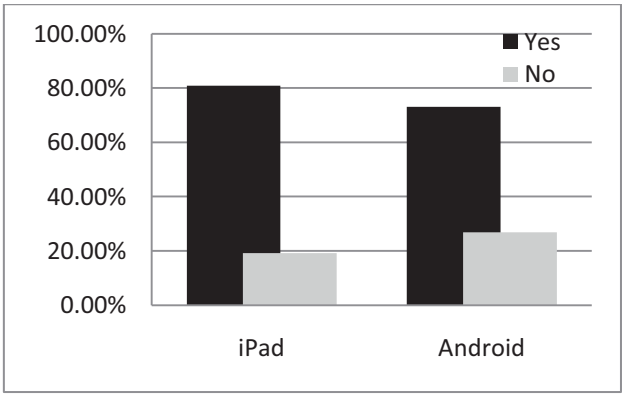

Figure 2. Survey results

This survey would have been followed up by a second survey if needed. However, the results provided a statistical basis for proposing the additional export functions for Sophie (see below).

Significance of the Results

Several factors are significant in evaluating these results:

First is the meaning of the response rate of 48%. Response rates vary widely depending on the type of relationship the survey provider has with potential respondents. The type of relationship of the survey provider to the respondents is an indicator of the level of response to be expected in relation to other surveys of the same type, and provides a measure of response quality. This survey was of type "no relationship". This is the most difficult situation – obtaining responses from those with whom the survey provider has no relationship. No relationship surveys typically have a response of 0% to 20%. The population in the iPad-Android survey was of the "no relationship" type since the members were

not from any unifying organization. There was no common bond between responders other than working in the same field at an upper level. The response rate of over 48% is significantly higher than the usually expected upper bound for “no relationship” marketing surveys.

Second is the quality of the respondent sample. All members of the sample were of a level to speak knowledgeably about e-book platforms in qualified companies and institutions. It was a random sample within the bounds of the expert population.

Third is the large skewing of the expressed opinion. The respondents show interest in the iPad and Android by responding at a 48% rate and by indicating the desirability of the iPad capability of 80% and Android capability of 73%, both quite high.

Fourth is the relationship of the sample size to the population size. Research of senior executives in September 2010 by the Frankfurt Book Fair involving all large and midsize publishers in the United States and the United Kingdom indicated that the population of such individuals is approximately 1500. These two countries are among those with highest level of adoption of digital publishing products worldwide. Based on this number, a goal of this survey was to obtain sufficient response to represent the views of a population of 5,000 such experts with a 90% confidence.

Statistical Significance

The positive response of 80.8% to the iPad question for the sample indicates a 90% confidence that a population of 5,000 individuals with the same level of familiarity with electronic publishing would respond positively between 70% and 95% concerning the importance of the iPad in their organizations.

The positive response of 73.1% to the Android question for the sample indicates a 90% confidence that a population of 5,000 individuals with the same level of familiarity with electronic publishing would respond positively between 58% and 88% concerning the importance of Android in their organizations.

Even the lower bound of extrapolated iPad responses shows a clear majority. The Android, while a smaller percentage, is still a majority. These results are taken as a good indicator that iPad and Android are considered to be important to publishers because they are important to readers, and by extension they are important to the spread of Sophie as an authoring environment. The intention of this survey was to provide a quick assessment of the merit of adding iPad, Android devices, or both as available platforms for Sophie books. The results give good support for adding both the iPad and Android as reading platforms.

Choosing a Format for Interactive Books. Drawbacks of EPUB

The EPUB format's purpose is to define the mechanisms by which the various components of a publication are tied together and to provide additional structure and semantics to the electronic publication. The “carrier” of the digital content is based on XHTML 1.1, resulting in a number of additional limitations over certain elements. For example, XHTML 1.1 does not support audio and video elements. Although many modern EPUB Readers (such as iBooks) do support HTML5 media elements in the XHTML book content, EPUB books are, by design, static. EPUB does not provide the means to define and incorporate interactivity elements in an e-book, which are important features of Sophie books.

HTML5 adds new features to HTML. Video, audio and “canvas” elements, as well as scalable vector graphics, have been added. Perhaps most significantly in the context of Sophie, however, is the improved handling of media-rich material on the Web, without having to deploy proprietary plug-ins and their APIs. Using HTML5 as a content carrier for electronic books would enable the use of multimedia and interactivity, the lack of which is a significant drawback of EPUB as an e-book format. Eventually, new versions of EPUB will potentially be based on new versions of XHTML and will support multimedia in the future. However, the unsupported content types are not the only limitation that prevents the use of the format for rich interactive books like the books created with Sophie. These limitations are additionally due to the underlying concepts of EPUB being generally a packaging format for XHTML content.

HTML5 is a young and evolving format but is being eagerly embraced by the Internet community. Even though HTML5 is currently a W3C working draft published in 2008, all major browser developers have already implemented support for its multimedia elements. Large industry leaders, such as Apple, Google, and Microsoft, display strong support for the technology and, through groups like WHATWG (http://www.whatwg.org/), drive it toward quick development and adoption. The current level of support for HTML5 by the WebKit-based Safari browser on the iPad and Android Web Browser enables the development of the e-book features enabled by Sophie. There is little doubt that support of these features will be continued in any future versions of HTML5, as they are at the base of the ideas that drive HTML5.

A recent article (Weiss, 2010) published by the American Society for Training & Development lists an overview of the coming benefits of HTML5 in e-learning:

- HTML5 improves video and audio quality extensively. As a result, your social learning/social media experience has just increased infinitely. Vivemo, Blip.tv, and YouTube (beta) are already offering videos in HTML5 players.
- HTML5 is open source.
- HTML5 provides enhanced multimedia.
- HTML5 works in the iPad tablet, and will work in other manufacturers' tablets and smart phones. M-learning to the next level, and e-Readers down the pipeline.
- HTML5 eliminates plug-ins.
- You can build courses in HTML5, as you would in Flash. As rapid e-learning authoring tools see HTML5 as commonplace, they will need to adapt, thus enabling HTML5 files to be inserted.
- HTML5 will offer mobile learning features that will far exceed anything that is possible today. (Weiss, 2010)

It is important to point out that HTML5 is not an e-book format. It is a markup language for defining rich and interactive Web pages which, together with JavaScript, provides all necessary technological tools to implement an e-book. The breadth of HTML5 makes it inconvenient as a book format since it does not impose any strict structure over the content. Page structure, metadata, interactive objects, and relations between book objects are pieces of book data that need to be presented in a structured, declarative, and well-defined way to enable consistent interpretation by various types of reader software (not necessarily web-based).

Therefore, although HTML5 provides the necessary means to build web-based readers which display dynamic content, it is not itself a convenient format for this content. The project team is currently working on a specification draft for a JSON-based declarative e-book format that reuses concepts of EPUB but allows for a well-structured definition of rich, interactive multimedia books. The results from this effort will be presented in a subsequent paper.

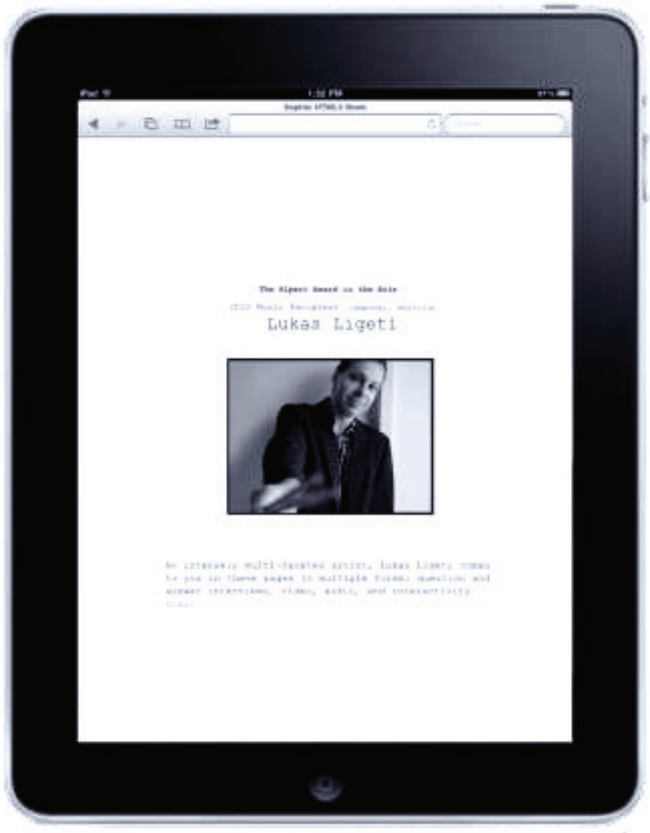

Figure 3. A view of one of the prototype Sophie 2.0 iPad books

Conclusion

Using HTML5 to make Sophie projects readable on a new generation of mobile devices and tablets makes complete sense and continues Sophie's ongoing evolution in dialogue with the culture and communities that both desire and need it. Content authoring platforms of today are facing the unavoidable challenge of supporting multi-channel publishing to the diverse types of widely used reading devices. For an e-book to reach its maximum audience, it needs to be accessible on all these devices in a consistent way. Among the available technology choices for distributing rich, interactive, electronic books for the Web, iPad and Android devices, HTML5 stands out as virtually the only feasible solution. Because of its support by the WebKit-based iOS and Android tablets and by an ever increasing number of desktop-based browsers, HTML5-based e-Readers will enjoy an increasing opportunity to provide a platform independent, consistent reading experience to all users, regardless of the hardware they are using.

Results of an industry survey showed that publishers believe that both iPad and Android are important to their planning. In addition, authors currently working with Sophie as part of the project to create six model books have expressed a desire to have their work available on the iPad. A feasibility study determined that the use of HTML5 as an output format would allow Sophie books to be read on both iPad and Android devices. While HTML5 is at the beginning of its life cycle, it clearly offers extended usefulness to Sophie. Indeed, using HTML5 would not only provide iPad and Android support, but would also support future HTML5 readers and attract additional developers from the open source community. The Sophie support team is currently focusing efforts on developing this capacity in tandem with the Exemplary Book project, with the goal of keeping Sophie relevant as the software continues to build a base of authoring knowledge.

Development began with a four-month feasibility study that reviewed EPUB and HTML5 as possible technologies for making Sophie books available on the iPad and Android devices. EPUB was chosen among the available e-book formats for inclusion in the study because of its widespread use and the

number of popular applications currently using it, including the iBooks and Nook Readers (Wikipedia, 2011).

References

Davidson, C.N. & Goldberg, D.T. (2009). *The future of learning institutions in a digital age*. Boston, MA: MIT Press. Retrieved May 29, 2011 from http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=11841

Drucker, J. (2009). BlindSpots. *The Chronicle of Higher Education*. April 3, 2009. Retrieved June 2010 from http://chronicle.com/article/Blind-Spots/9348/

Gartner. (2010). *Gartner says android to become No. 2 worldwide mobile operating system in 2010 and challenge symbian for No. 1 position by 2014*. Retrieved January 9, 2011 from http://www.gartner.com/it/page.jsp?id=1434613

Georgiev, K. & Sredkov, M. (2010). Sophie 2.0 - A platform for reading and writing of interactive multimedia books in a networked environment. In T. Hedlund and Y. Tonta (Eds.) *Publishing in the Networked World: Transforming the Nature of Communication. 14th International Conference on Electronic Publishing 16-18 June 2010, Helsinki, Finland* (pp. 437-443). Helsinki: Hanken School of Economics.

Johnson, L., Smith, R., Levine, A., & Haywood, K. (2011). *The 2011 Horizon Report*. Austin, Texas: The New Media Consortium. Retrieved May 29, 2011 from http://www.educause.edu/Resources/2011HorizonReport/223122.

Weiss, C. (2010). HTML5: Game changer for e-learning? *American Society for Training & Development*. Retrieved January 15, 2010, from http://www.astd.org/LC/2010/0910_weiss.htm

Wikipedia (2011). *Comparison of e-book formats*. Retrieved January 24, 2011 from hhttp://en.wikipedia.org/wiki/Comparison_of_e-book_formats.

The Impact of Open Access Contributions: Developed and Developing World Perspectives

Sheikh Mohammad Shafi

Department of Library & Information Science, University of Kashmir. Srinagar (J&K) India-190006. E-mail: smshafi@kashmiruniversity.ac.in

Mohammad Haneef Bhat

Islamia College of Science &Commerce. Srinagar (J&K) India-190002. E-mail: mhanief30@yahoo.co.in

Abstract: *The study explores the research impact of 'Open Access research articles' across the globe with a view to test the hypothesis that "OA research contributions emanating from developing countries receive equal citations (subsequently resultant research impact) as those from the developed world". The study covers 5639 research articles from 50 Open Access DOAJ based Medical Sciences journals covering the period from 2005 to 2006. The research impact of OA research publications measured by the citation counts varies from journal to journal and from country to country. Statistically significant difference is noted between the research impact of the developed and the developing world for OA research articles. The research articles from the developed countries receive higher number of citations (subsequently resultant research impact) compared to those of the developing world. The study may help and pave way for framing policies and strategies to increase the impact of research in the developing world.*

Keywords: *Open Access; research impact; DOAJ.*

Introduction

Open Access (OA) to scientific literature means removal of barriers (including price barriers) from accessing scholarly work (Eysenbach, 2006). Open Access to scholarly literature has received increasing attention in the academic, research and publishing circles over the last decade and is a hot topic for deliberations and discussions. Many scholars, organizations and institutions across the globe are working towards the cause, as a result of which more and more of the scholarly literature is becoming openly accessible. The growth of OA movement is particularly in response to the enormous costs of many scholarly journals. With traditional journal publication methods, it is not uncommon for an institution to have to pay for an article twice. First they pay scholars to produce the work and then the institution pays to purchase the work back from the journal publisher (Corrado, 2005). In this situation, the open access movement gains worldwide support as an alternative and sustainable model of scholarly communications and accessing research literature. The number of OA journals publishing high quality, peer-reviewed research is growing. SPARC and SPARC Europe are in partnership with a number of these journals (Prosser, 2004). This is evident from DOAJ (Directory of Open Access Journals), listing fully peer-reviewed OA journals when launched in May 2003 with 375 items, now figure over with 6271 in 2011.

Giving all interested readers access accelerates research, enriches education, sharing learning among rich and poor nations and ultimately enhance return on investment in research (much of which comes from the world's tax payers) (Prosser, 2004). The developed world consists of information rich countries, enterprises and organizations. These enterprises control over the valuable information resources. The developing world is at the critical junctions where the development of technologies, economics and humanity largely depends on access to relevant and adequate information resources. The developing world consist of information poor least-developed countries, where the research institutions cannot afford to subscribe a wide array of primary literature due to resources crunch or limited budgetary provisions. As such, for the developing world open access movement has come as a boon (Ghosh & Das, 2006). Hence, the present study endeavors to make a comparative study of

research impact of OA articles between the developed and the developing world through citation analysis.

Review of Literature

The failure of traditional scholarly communication model led to the development of institutional repositories (IRs) and OA journals (Prosser, 2003), as they can better serve the international research community by providing a fairer, more equitable and more efficient system of scholarly communication (Prosser, 2004). The OA journals and OA repositories are steps towards democratization of information and knowledge as they remove access restrictions (Ylotis, 2005). It is due to these benefits that OA is gaining momentum with support from library and professional groups, university faculties and even journal publishers (Falk, 2004), and is bringing great benefits to the academic world (McCulloch, 2006).

Open Access enriches the global knowledge base by incorporating the missing research from the developing world and improves the global knowledge flow (Chan & Costa, 2005), and at the same time it connects the developing world to the system of science by providing access to scientific literature published in the developed world (Haider, 2007). As a result, developing countries have embraced OA with a view to promoting the visibility of research done in these regions (Fernandez, 2006). The majority of authors in the developing world are not well informed on how they could enhance the visibility of their publications (Ramachandran & Scaria, 2004). However with the passage of time the proportion of authors publishing in OA journals has grown from 11% in 2004 to 29% in 2005 (Rowlands & Nicholas, 2005). As it is now well established that OA makes research papers more visible and increases their research impact (Lawrence, 2001; Antelman, 2004; Harnad & Brody, 2004; Hajjem, Harnad, & Gingras, 2005; Eysenbach, 2006; Hajjem, Gingras, Brody, Carr & Harnad, 2005), although with variations across disciplines (Tonta, Unal, & Al, 2007; Shafi, 2008; Shafi & Bhat, 2008).

Problem

The research carried out in developing countries representing 80 percent of the world's population is largely invisible to international science because of economic constraints. It is believed that the low research impact of research output of the developing countries is due to its poor visibility (Fernandez, 2006). The OA has overcome that barrier between the developing and the developed world to a large extent. However, it remains to be seen whether the OA research contributions from the developing world receive the same amount of citations (and the resultant research impact) as those from the developed world especially during the recent decade.

Objectives

The following objectives are put forth for the study:

- To assess the research impact of OA journal articles across countries;
- To compare the research impact of OA journal articles across developed and developing world;
- To verify the hypothesis drawn.

Scope

The scope of the study is limited to Open Access articles appearing in English language 'OA' journals in the field of Medical Sciences, as it has been found that Open Access journals are much popular in the field of Medical Sciences, in comparison to other fields (Bhat, 2008; Bhat, 2009).

Methodology

The directory of Open Access journals (www.doaj.org) is used to select OA journals in the field of Medical Sciences. DOAJ listed 416 journals in this field (January, 2011). Mono-lingual (English

language) journals having back issues available since 2005 reduced the number to 142 journals (93 from the developed world and 49 from the developing world). 25 journals from each group were selected randomly using function "sample" of 'R' software (R Development Core Team, 2011). The classification of countries into developed and developing economies is taken from UN Human Development Index and IMF report.

The total number of articles published in OA journals in 2005 and 2006 were identified for all the 50 journals. The articles of each journal were arranged in chronological order, and a sample of 10 articles (five each from year 2005 and 2006) is randomly selected using function "sample" of 'R' software (R Development Core Team, 2011), thereby making a total of 500 articles (250 articles from each group).

With a view to show the evidence of broader international research impact beyond the indexed journals, Google Scholar is used for finding the citations instead of WOS and Scopus databases (although all the 50 journals covered by the study were found indexed in Scopus and some in WOS). It has been found that Google Scholar provided the largest citations compared to WOS and Scopus for recent articles of Journal of American Society for Information Science and Technology (Baur & Bakkalbasi, 2005). Also Google Scholar provided the highest number of unique citations for current oncology articles (Bakkalbasi, Baur, Glover, & Wang, 2006).

All the 500 articles were searched in Google Scholar for citations (March, 2011). The number of citations were recorded for each article (Table 1). The data is tabulated and analysed in a systematic order to reveal findings in accordance with the laid down objectives. Different standard statistical techniques are used to estimate various statistical tests and verification of hypothesis.

Table 1. Sampling statistics

S. No.	Country Group	No. of journals in DOAJ (having back issues since 2005)	No. of OA journals in sample	No. of articles			No. of articles taken for sample
				2005	2006	Total	
1.	Developed	93	25	888	1318	2206	250 (11.3)
2.	Developing	49	25	1633	1800	3433	250 (7.3)
	Total	142	50	2521	3118	5639	500 (8.9)

(The numbers in parentheses indicate percentage)

Hypothesis

The present study focuses to test the hypothesis "OA research contributions emanating from developing countries receive equal citations (subsequently resultant research impact) as those from the developed world".

Results and Discussion

The total number of citations received by 500 Open Access research articles is 4543 (Mean=9.08, S.D. =14.95). The distribution of citations among the research articles is somewhat skewed as the Standard Deviation is greater than Mean (Table 2). Out of 500 research articles, 124 received no citations and 54 did one citation each. Out of 4543 citations, 1131 (24.89%) are recorded by just 20 OA research publications (Table 5 & 6).

Table 2. Citation count of OA articles

S. No.	Country Group	No. of articles			Sample	Citations	Mean	S.D.
		2005	2006	Total				
1.	Developed	888	1318	2206	250	3344	13.37	18.25
2.	Developing	1633	1800	3433	250	1199	4.79	8.81
	Total	2521	3118	5639	500	4543	9.08	14.95

The 250 research papers from 25 OA journals of the developed world received 3344 citations (Mean=13.37, S.D.=18.25). The average number of citations varies from 0.1 for *Acta Medica Nagasakiensia (Japan)* to 29.5 for *BMC Medical Informatics and Decision Making (UK)*. The standard deviation also varies from journal to journal with *Acta Medica Nagasakiensia (Japan)* having S.D. of 0.31 on one hand and *BMC Medical Research Methodology (UK*) having S.D. of 38.03 on the other hand. Out of 25 journals, 14 journals have S.D. greater than Mean. Forty-six (out of 250) articles received no citations and 8 articles received one citation each (Table 3). Ten research papers generated almost a quarter (22.51%, or 753 citations) of all citations (Table 5).

Table 3. Citation count of OA articles (Developed World Group)

S. No	Journal Title	Country	Number of articles						
			2005	2006	Total	*n*	Cit.	*X*	S.D.
1.	Acta Medica Nagasakiensia	Japan	30	35	65	10	1	0.1	0.31
2.	Archives of Medical Science	Poland	31	39	70	10	20	2.0	2.49
3.	Biomedical Research	Japan	40	39	79	10	8	0.8	1.47
4.	BMC Complementary and Alternative Medicine	UK	21	41	62	10	143	14.3	8.06
5.	BMC Medical Genetics	UK	43	88	131	10	203	20.3	21.58
6.	BMC Medical Informatics and Decision Making	UK	38	40	78	10	295	29.5	24.41
7.	BMC Medical Research Methodology	UK	37	61	98	10	246	24.6	38.02
8.	Cell Communication and Signaling	UK	14	8	22	10	73	7.3	5.86
9.	Communicable diseases intelligence	Australia	53	51	104	10	44	4.4	4.92
10.	Head & Face Medicine	UK	14	48	62	10	59	5.9	5.25
11.	International Journal of Behavioral Nutrition and Physical Activity	UK	16	44	60	10	276	27.6	19.62
12.	Journal of Applied Clinical Medical Physics	USA	41	38	79	10	59	5.9	8.15
13.	Journal of Biomedicine and Biotechnology	USA	38	22	60	10	174	17.4	18.66
14.	Journal of Circadian Rhythms	UK	15	16	31	10	38	3.8	3.04
15.	Journal of International Medical Research	UK	81	87	168	10	89	8.9	14.81
16.	Journal of Medical Internet Research	Canada	31	19	50	10	278	27.8	16.62
17.	Journal of Medical Investigation	Japan	48	41	89	10	47	4.7	4.05
18.	Journal of Negative Results in BioMedicine	UK	12	19	31	10	87	8.7	10.43
19.	Journal of Neuroinflammation	UK	29	33	62	10	80	8.0	7.54
20.	Journal of the American Board of Family Medicine	USA	52	65	117	10	270	27.0	13.49
21.	Korean Journal of Radiology	S. Korea	46	47	93	10	153	15.3	22.49
22.	Molecular Pain	UK	36	38	74	10	220	22.0	21.91
23.	Morbidity and Mortality Weekly Report (MMWR)	USA	11	266	277	10	262	26.2	33.46
24.	Upsala Journal of Medical Sciences	Sweden	19	27	46	10	59	5.9	9.97
25.	Virology Journal	UK	92	106	198	10	160	16.0	12.00
	Total		888	1318	2206	250	3344	13.37	18.25

1199 citations are received by 250 articles of 25 OA journals from the developing world (Mean=4.79 & S.D. =8.81). The standard deviation varies from journal to journal (with 0.42 for the *Journal of Research in Medical Sciences, Iran* to 26.62 for the *Brazilian Journal of Medical and Biological Research, Brazil*), as also the average number of citations (with 15.6 for *Brazilian Journal of Medical and Biological Research, Brazil* to 0.2 for *Journal of Research in Medical Sciences, Iran*) (Table 4).

For the 15 journals (out of 25) the S.D. is higher than the Mean. Seventy-eight papers received no citations and 46 did one citation each. Ten research papers generated almost one third (31.52%, or 378 citations) of all citations (Table 6)

Table 4. Citation count of OA articles (Developing world group)

S. No	Journal Title	Country	No. of articles						
			2005	2006	Total	*n*	Cit.	*X*	S.D.
1.	Acta Medica Iranica	Iran	86	90	176	10	8	0.8	1.54
2.	African Health Sciences	Uganda	54	47	101	10	79	7.9	5.85
3.	African Journal of Health Sciences	Kenya	8	30	38	10	14	1.4	1.77
4.	Annals of African Medicine	Nigeria	34	42	76	10	12	1.2	1.13
5.	Biological Research	Chile	37	54	91	10	107	10.7	9.93
6.	Brazilian Journal of Medical and Biological Research	Brazil	200	170	370	10	156	15.6	26.62
7.	Chinese Medical Journal	China	197	217	414	10	84	8.4	7.91
8.	Clinics	Brazil	50	61	111	10	130	13.0	7.64
9.	Croatian Medical Journal	Croatia	96	76	172	10	111	11.1	14.53
10.	European Journal of General Medicine	Turkey	39	43	82	10	7	0.7	1.25
11.	Indian Journal of Medical Research	India	97	110	207	10	83	8.3	5.85
12.	Indian Journal of Medical Sciences	India	54	38	92	10	43	4.3	3.43
13.	Indian Journal of Palliative Care	India	20	8	28	10	25	2.5	3.77
14.	Iranian Biomedical Journal	Iran	32	33	65	10	18	1.8	2.20
15.	JK Science	India	45	41	86	10	9	0.9	0.99
16.	Journal of Medical Sciences	Pakistan	70	187	257	10	5	0.5	1.26
17.	Journal of Postgraduate Medicine	India	34	32	66	10	77	7.7	10.95
18.	Journal of Research in Medical Sciences	Iran	72	58	130	10	2	0.2	0.42
19.	Journal of the Association of Physicians of India	India	88	102	190	10	36	3.6	3.20
20.	Kathmandu University medical journal	Nepal	71	91	162	10	19	1.9	4.01
21.	Malaysian Journal of Medical Sciences	Malaysia	16	22	38	10	8	0.8	1.31
22.	Online Journal of Health and Allied Sciences	India	16	21	37	10	12	1.2	1.22
23.	Pakistan Journal of Medical Sciences	Pakistan	91	91	182	10	25	2.5	2.27
24.	Revista do Instituto de Medicina Tropical de Sao Paulo	Brazil	70	71	141	10	121	12.1	11.76
25.	Turkish Journal of Medical Sciences	Turkey	56	65	121	10	8	0.8	1.54
	Total		1633	1800	3433	250	1199	4.79	8.81

The research articles from developed countries have received almost three times more citations than those from the developing nations. The journals from USA and UK have received more citations than those from Japan, Poland, Australia, and Sweden within the developed nations group. A few journals from the developing world have also received double digit mean citations. These include journals from the countries like Chile, Brazil and Croatia. It is interesting to note that the lowest research impact journal in the study is from a developed country i.e., Japan (*Acta Medica Nagasakiensia*).

The research impact of OA research publications varies a great deal from journal to journal and from country to country within the group. Whereas OA journals from the developing world have emerged during the last decade, most of the OA journals from the developed world are well established, and have lately switched to Open Access mode with their clientele. The OA research papers from the developing countries despite lagging behind their counterparts' vis-à-vis research impact have shown

good performance (which would not have been possible without Open Access). With improvement in quality of research in developing countries (as a result of economic development and liberal allocation of funds for R & D activities), the situation is going to improve further in the future.

Table 5. Top ten most cited articles (Developed world group)

S. No	Article title	Country	Journal Title	Cit.
1.	Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies/Penny F Whiting, Marie E Weswood, et al./ V6/ 2006	UK	BMC Medical Research Methodology	126
2.	Surveillance for Dental Caries, Dental Sealants, Tooth Retention, Edentulism, and Enamel Fluorosis --- United States, 1988--1994 and 1999--2002/ 2005	USA	Morbidity and Mortality Weekly Report (MMWR)	106
3.	IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study/Elske Ammenwerth, Carola Iller, Cornelia Mahler/ V6/ 2006	UK	BMC Medical Informatics and Decision Making	73
4.	Cytogenetic abnormalities and fragile-x syndrome in Autism Spectrum Disorder/ Kavita S Reddy/ V6/ 2005	UK	BMC Medical Genetics	72
5.	False Positive and False Negative FDG-PET Scans in Various Thoracic Diseases /JM Chang, HJ Lee et al/ V7/ 2006	S. Korea	Korean Journal of Radiology	72
6.	Neuropathic pain develops normally in mice lacking both $Na_v1.7$ and $Na_v1.8$ /Mohammed A Nassar, Alessandra Levato, et al./ V1/ 2005	UK	Molecular Pain	71
7.	Dicer-Derived MicroRNAs Are Utilized by the Fragile X Mental Retardation Protein for Assembly on Target RNAs/ Isabelle Plante, Laetitia Davidovic, et al./ 2006	USA	Journal of Biomedicine and Biotechnology	66
8.	Extracting principal diagnosis, co-morbidity and smoking status for asthma research: evaluation of a natural language processing system/ Qing T Zeng, Sergey Goryachev, et al./ V6/ 2006	UK	BMC Medical Informatics and Decision Making	59
9.	Internet Interventions for Long-Term Conditions: Patient and Caregiver Quality Criteria/C Kerr, E Murray et al/ 7/ 2006	Canada	Journal of Medical Internet Research	55
10.	Pulmonary Hypertension Surveillance --- United States, 1980--2002/ 2005	USA	Morbidity and Mortality Weekly Report (MMWR)	53

Table 6. Top ten cited articles (Developing world group)

S. No	Article	Country	Journal Title	Cit.
1.	Geographic distribution of hepatitis C virus genotypes in Brazil/ Campiotto, S.; Pinho, J.R.R.; et al./ V38/ 2005	Brazil	Brazilian Journal of Medical and Biological Research	89
2.	Challenges of DNA Profiling in Mass Disaster Investigations/ Antonio Alonso, Pablo Martín, et al. /V46/ 2005	Croatia	Croatian Medical Journal	46
3.	Conversion from laparoscopic to open cholecystectomy: Multivariate analysis of preoperative risk factors/ M Tayeb, SA Raza et al./ V51/ 2005	India	Journal of Postgraduate Medicine	38
4.	Posaconazole treatment of refractory eumycetoma and chromoblastomycosis/ Negroni, Ricardo; Tobón, Angela; et al /V47/ 2005	Brazil	Revista do Instituto de Medicina Tropical de Sao Paulo	35
5.	A Rapid and Efficient Method for Purifying High Quality Total RNA from Peaches (Prunus persica) for Functional Genomics Analyses/ Meisel, Lee; Fonseca, Beatriz; et al/ V38/ 2005	Chile	Biological Research	32
6.	Evolutive standard base excess and serum lactate level in severe sepsis and septic shock patients resuscitated with early goal-directed therapy: still outcome markers?/ Park, Marcelo; Azevedo, Luciano Cesar Pontes; et al./ V61/ 2006	Brazil	Clinics	30
7.	Depressive Symptomatology among University Students in Denizli, Turkey: Prevalence and Sociodemographic Correlates/ Mehmet Bostanci, Osman Ozdel, et al./ V46/ 2005	Croatia	Croatian Medical Journal	28
8.	Development and potential of a biomimetic chitosan/type ¢ò collagen scaffold for cartilage tissue engineering / SHI De-hai, CAI Dao-zhang, et al./ V118/ 2005	China	Chinese Medical Journal	28
9.	Blastocystis hominis and other intestinal parasites in a community of Pitanga City, Paraná State, Brazil/ Nascimento, Solange Aparecida; Moitinho, Maria da Luz Ribeiro/ V47/ 2005	Brazil	Revista do Instituto de Medicina Tropical de Sao Paulo	27
10.	Molecular and pathological basis of aceruloplasminemia/ Kono, Satoshi; Miyajima, Hiroaki/ V39/ 2006	Chile	Biological Research	25

Testing Hypothesis

The Chi square (χ^2) test is carried out to test the null hypothesis "OA research contributions emanating from developing countries receive the equal citations (subsequently resultant research impact) as those from the developed world".

$$\chi^2 = \Sigma (O-E)^2/E$$

E = Expected citations

O = Observed citations

The value of χ^2 is 1012.73. At 1 df (degree of freedom) this is highly significant at .05 level of significance. Thus the null hypothesis "OA research contributions emanating from developing countries receive equal citations (subsequently resultant research impact) as those from the developed world" is rejected and it is concluded that OA research contributions emanating from developing countries receive fewer citations (subsequently resultant research impact) as those from the developed world.

Conclusion

The research impact of Open Access research articles varies from country to country and from journal to journal within countries. The research papers from the developed countries receive more citations (and the resultant research impact) compared to research publications from the developing countries.

Since the present study is limited to OA research publications in the field of Medical Science, there is a need to verify the trend on a large canvas covering different disciplines, including Social Sciences. As the visibility problem has no relevance in the Open Access environment, the underlying factors for low research impact of research papers from the developing world need to be ascertained and rectified. The reasons could be many like the quality of research, peer review system, etc. It may also be due to the fact that a large number of authors from the developing countries still prefer well established journals from the developed world for disseminating their research output. Besides Open Access, rigorous peer review standards will have to be enforced by the journals emanating from the developing world to ensure maximal research impact of their publications at par with their counterparts in the developed world.

References

Antelman, K. (2004). Do open-access articles have a greater research impact? *College & Research Libraries*, 65(5). Retrieved December 28, 2010 from http://eprints.rclis.org/archive/00002309/01/do_open_access_CRL.pdf

Bakkalbasi, N., Baur, K., Glover, J. & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical Digital Libraries*, 7. Retrieved December 29, 2010 from http://www.bio-diglib.com/content/pdf/1742-5581-3-7.pdf.

Baur, K. & Bakkalbasi, N. (2005). An examination of citation counts in a new scholarly communication environment. *D-Lib Magazine*, 11(9). Retrieved December 29, 2010 from http://www.dlib.org//dlib/september05/bauer/09bauer.html

Bhat, M. H. (2008). Open access publishing in Jammu and Kashmir State: an assessment. *Trends in Information Management (TRIM),* 4(1), 20-37

Bhat, M. H. (2009). Open access publishing in Indian premier research institutions. *Information Research*, 14(3). Retrieved December 20, 2010 from http://informationr.net/ir/14-3/paper409.html

Chan, L., & Costa, S. (2005). Participation in the global knowledge commons: challenges and opportunities for research dissemination in developing countries. *New Library World*, 106 (3/4). Retrieved December 20, 2010 from http://www.emeraldinsight.com/10.1108/03074800510587354

Corrado, E.M. (2005). The importance of open access, open source, and open standards for libraries. *Issues in Science and Technology Librarianship*. Retrieved December 20, 2010 from, http://www.istl.org/05-spring/article2.html

Eysenbach, G. (2006). Citation advantage of open access articles. *PLOS Biology*, 4(5). Retrieved December 29, 2010 from http://biology.plosjournals.org/archive/1545-7885/4/5/pdf/10.1371_journal. pbio. 0040157-L.pdf

Falk, H. (2004).Open access gains momentum. *The Electronic Library*, 22(6). Retrieved December 20, 2010 from http://www.emeraldinsight.com/10.1108/02640470410570848

Fernandez, L. (2006). Open access initiatives in India- an evaluation. *The Canadian Journal of Library and Information Practice and Research*, 1(1). Retrieved December 21, 2010 from http://journal.lib.uoguelph.ca/index.php/perj/article/view/110/172

Ghosh, S. B., & Das, A. K. (2006). Open access and institutional repositories – a developing country perspective: a case study of India. World Library and Information Congress:72nd IFLA General Conference and Council, August 20-24, 2006, Seoul, Korea. Retrieved January 5, 2011 from https://drtc.isibang.ac.in/handle/1849/272

Haider, J. (2007). Of the rich and the poor and other curious minds: on open access and "development". *Aslib Proceedings*, 59(4/5). Retrieved December 22, 2010 from http://www.emeraldinsight.com/10.1108/00012530710817636

Hajjem, C., Harnad, S., & Gingras, Y. (2005). Ten-year Cross-disciplinary comparison of the growth of open access and how it increases research citation impact. Retrieved December 20, 2010 from http://arxiv.org/ftp/cs/papers/0606/0606079.pdf

Hajjem, C., Gingras, Y., Brody, T., Carr, L., & Harnad, S. (2005). Open access to research increases citation impact. Retrieved December 24, 2010 from http://eprints.ecs.soton.ac.uk/11687/

Harnad, S., & Brody, T. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. *D-lib Magazine*, 10(6). Retrieved December 20, 2010 from http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/june04/harnad/06harnad.html

Lawrence, S. (2001). Online or invisible? *Nature*, 411(6837). Retrieved December 24, 2010 from http://citeseer.ist.psu.edu/cache/papers/cs/29685/http:zSzzSzwww.ipo.tue.nlzSzhomepageszSzmrauterbzSzpublicationszSzCITESEER2001online-nature.pdf/lawrence01online.pdf

McCulloch, E. (2006). Taking stock of open access: Progress and issues. *Library Review*, 55(6). Retrieved December 20, 2010 from http://eprints.cdlr.strath.ac.uk/2325/01/Mcculloch_openaccessELIS.pdf

Prosser, D. (2003). Institutional repositories and open access: the future of scholarly communication. *Information Services & Use*, 23(2-3). Retrieved December 22, 2010 from http://iospress.metapress.com/media/3duf6utqrk4vyp1kvrt0/contributions/b/0/t/9/b0t9y69afjhylw9y.pdf

Prosser, D.C. (2004). The next information revolution-how open access repositories and journals will transform scholarly communications. *LIBER Quarterly: The Journal of European Research*, 14(1-4). Retrieved December 20, 2010 from http://webdoc.sub.gwdg.de/edoc/aw/liber/lq-1-04/prosser.pdf

R Development Core Team. (2011). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved December 28, 2010, from http://www.R-project.org

Ramachandran, P.V., & Scaria, V. (2004). Open access publishing in the developing world: making a difference. *Journal of Orthopaedics,* 1(1). Retrieved December 20, 2010 from http://jortho.org/2004/1/1/e1/index.htm

Rowlands, I., & Nicholas, D. (2005). Scholarly communication in the digital environment: The 2005 survey of journal author behaviour and attitudes. *Aslib Proceedings*, 57(6). Retrieved December 20, 2010 from http://www.emeraldinsight.com/10.1108/00012530510634226

Shafi, S.M. (2008). Research impact of open access research contributions across disciplines. *Proceedings ELPUB 2008 Conference in Electronic Publishing, Toronto, Canada*. Retrieved December 28, 2010 from http://elpub.scix.net/data/works/att/343_elpub2008.content.pdf

Shafi, S.M., & Bhat, M. H. (2008). Impact of open access research contributions across select disciplines. *Trends in Information Management (TRIM)*, 4(1), 108-126

Tonta, Y., Unal, Y., & Al, U. (2007). The research impact of open access journal articles. Edited by: Leslie Chan and Bob Martens (Eds.). *ELPUB2007. Openness in Digital Publishing: Awareness, Discovery and Access - Proceedings of the 11th International Conference on Electronic Publishing held in Vienna, Austria 13-15 June 2007* (pp. 321-330). Retrieved December 28, 2010 from, http://eprints.rclis.org/archive/00009619/01/tonta-unal-al-elpub2007.pdf

Ylotis, K. (2005). The open access initiative: a new paradigm for Scholarly Communications. *Information Technology and Libraries*, 24(4). Retrieved December 18, 2010 from http://find.galegroup.com/itx/start.do?prodId=ITOF&userGroupName=bcdelhi

Institutional Repositories in Irish Third Level Institutions: Collaborating for Visibility, Reusability and Access[1]

Jerald Cavanagh

Limerick Institute of Technology, Moylish Park, Limerick, Republic of Ireland.
Email: Jerald.Cavanagh@lit.ie

Padraig Kirby

Limerick Institute of Technology, Moylish Park, Limerick, Republic of Ireland.
Email: Padraig.Kirby@lit.ie

Abstract: *This paper will describe the development of Institutional Repositories (IRs) in Irish third level institutions – starting with the development of stand-alone repositories in the mid 2000's and the move later in that decade towards collaboration and resources sharing through the development of National Research Portals and also Learning Resource Portals. We will briefly look at the various institutional repositories in existence in Ireland and also National Portals such as RIAN which are helping to make Irish research more visible and accessible. Some important European Union initiatives will also be dealt with and finally we will look briefly at the NDLR – Ireland's National Digital Learning Repository.*
Keywords: *Institutional repositories; research; collaboration.*

Introduction

At the most basic and fundamental level, an institutional repository is a recognition that the intellectual life and scholarship of our universities will increasingly be represented, documented, and shared in digital form, and that a primary responsibility of our universities is to exercise stewardship over these riches: both to make them available and to preserve them. An institutional repository is the means by which our universities will address this responsibility both to the members of their communities and to the public. It is a new channel for structuring the university's contribution to the broader world, and as such invites policy and cultural reassessment of this relationship (Lynch, 2003).

IRs have been in existence in libraries since the mid 2000s (Lawton, 2009). Emerging initially in the academic library sector in large institutions such as Massachusetts Institute of Technology and the University of California, over the last decade Irish institutions have been developing IRs as a means of preserving and promoting their research output across the globe. In the Irish context the National University of Ireland (NUI) Maynooth were pioneers in this field and have the distinction of being the first Irish university to host an institutional repository with their ePrints initiative launched in 2003. In 2006 TARA (Trinity Access to Research Archive) was launched and this was followed by University College Dublin's Research Online in 2007. In 2010 all Irish universities and some other third level institutions such as Institutes of Technology had institutional repositories. However in recent years the focus has shifted more towards collaboration, pooling and sharing of resources. While up to recently Irish universities were maintaining IRs independently, important developments have been the IReL-Open project - an Irish Universities Association (IUA) initiative which is building online open access institutional repositories in all seven Irish universities and linking these together through a National Research Portal. A very significant result of this project is the RIAN National Research Portal which went live in June 2010 and is available at http://www.rian.ie/. For the first time, Irish research will be made freely available on a worldwide scale. The development of IRs has

[1] An earlier version of this work appeared in Cavanagh, J. & Kirby, P. (2010) Institutional repositories in Irish third level institutions: collaborating for visibility, reusability and access*, Romanian Review of Library and Information Science,* No.3, pp. 41-50. Retrieved May 31, 2011 from http://www.abr.org.ro/www.abr.org.ro/rrbsi/rrbsi3_2010_en_sum.pdf.

not just focused on research. The development of IRs for collaboration and a community for sharing digital teaching and learning resources has seen widespread collaboration amongst third level institutions for the creation of Ireland's National Digital Learning Repository (NDLR) enabling sharing, collaboration and the development of learning resources (NDLR, 2010).

We will now describe the development of IRs in Irish universities − starting with the development of stand-alone repositories in the mid to late 2000s and the move towards collaboration and resource sharing through National Portals developed through initiatives like IReL-Open. We will also briefly look at some European Union initiatives which are helping to make Irish research more visible, and finally we will look at the NDLR − Ireland's National Digital Learning Repository.

A Review of Irish Institutional Repositories

As noted by Brennan (2007), Open Access Repositories are being constructed against the backdrop of heightened competition within the international research community and increased investment in research by governments, funding councils and the European Commission. Through initiatives such as the Programme for Research in Third-Level Institutions (PRTLI), the Strategy for Science, Technology and Innovation (SSTI) 2006-2013 and other forward looking initiatives there has been an enormous boost to funding across various disciplines such as arts, humanities and social sciences. Increased investment and competition leads to growing demands for accountability and evaluation of publicly funded research, and the failure of any Irish University to feature in the 'Top 100' of the 2005 Shanghai Jiao Tong University's Academic Ranking of World Universities caused us to raise "concerns about the quality of our universities and how we might deliver on our aspirations" (Browne, 2005 cited in Brennan, 2007). However some commentators have highlighted flaws in the Shanghai Ranking noting that it "credits a university with citations to publications which were written when an author was employed and supported by a rival university" (Brennan, 2007).

All Irish Universities and some Institutes of Technology in Ireland have IRs. Different content strategies have been adopted in developing their IRs. Several include depth and content while others emphasise breadth and IR size. We will now briefly examine these IRs:

ARAN (Access to Research at National University of Ireland, Galway http://aran.library.nuigalway.ie/xmlui/) is NUI Galway's institutional repository. ARAN collects, preserves and makes freely available scholarly communications, including peer-reviewed articles, working papers and conference papers created by NUI Galway researchers. Where material has already been published it is made available subject to the open-access policies of the original publishers. Like many other universities it promotes Open Access through Google, promotes the role of Open Access generally as well as initiatives such as Sherpa Juliet for funding. It advocates Sherpa Romeo for publishing and journal policies, stresses a commitment to the long term preservation and archiving of material deposited in the system and provides download statistics. It uses DSPACE, is multidisciplinary and users are encouraged to deposit journal articles, book chapters, working papers as well as conference papers. Material in ARAN is stored in collections based on the organisational structure of the University. At the highest level, collections in ARAN are divided up by College. ARAN then follows the organisational hierarchy of the University down to discipline. Within each discipline items are arranged by type, for example 'scholarly articles', 'conference papers' or 'working papers'. There are also collections for research institutes and centres. It is also possible to map an item from one collection to another, so a single item can appear in more than one collection if a member of staff has multiple affiliations. ARAN is one of the smaller IRs and up to 12-5-2009 ARAN had 62 items deposited (OPENDOAR, 2010).

Arrow@DIT http://arrow.dit.ie/ is the institutional repository of Dublin Institute of Technology which provides research output of the institution and up to 16-4-2010 Arrow had 1577 items on deposit. Content is chiefly articles, book chapters, conference papers, theses and reports with a useful browse function which allows publications to be sourced by Faculty/Department. All items in this collection are full text. This collection is being developed and is not currently a full listing. The term *publication* is used in the broadest sense as to encompass all forms of scholarly communication. It includes other useful features such as top 10 downloads, recent additions and a "paper of the day" alert. It uses digital commons software, users may register for RSS feeds to be alerted to new content,

and it encourages users to check the Sherpa Romeo database to find the publisher's policy with regard to self archiving.

CORA (Cork Open Research Archive) http://cora.ucc.ie/ is the institutional repository of University College Cork, Ireland. CORA provides access to scholarly research including peer-reviewed journal articles and electronic theses from across the four colleges and research institutes/centers. Users may set up RSS feeds to be alerted to new content. It uses DSPACE, is multidisciplinary, and is comprised of collections of peer-reviewed journal articles, research papers, book chapters, and other research materials as well as UCC research theses. Like ARAN this is still in development with just 24 deposited items up to 22-6-2009 (OPENDOAR, 2010).

DORAS (DCU Online Research Access Service) http://www.doras.dcu.ie is an online open access repository of research papers from Dublin City University and contains "a growing number of journal articles, conference papers, books, book chapters and theses from DCU authors" (Hill, 2008). Most of these papers are open access, and full-text papers are also indexed by search engines, thus making them easily searchable and retrievable using Google, Yahoo and Google Scholar. As noted by Hill (2008) nowadays "the only barrier that stands between an author and making his/her research open access is the time and effort required to upload the paper to the institutional repository". In the case of DORAS, DCU Library uploads the papers on behalf of the author − all the author has to do is email the file(s) to the Library, which will then check the paper for copyright clearance and upload the file(s) and bibliographic information associated with the paper. DORAS uses search engines effectively and in the first half of 2008, for example, 68% of traffic to the DORAS website came via search engines (Hill, 2008). It is structured in a web-friendly manner that makes it easy for search engines to crawl and index the web pages and full text files. Two-thirds of users accessing DORAS through search engines access it through Google, and Hill (2008) has emphasised the high visibility of DORAS papers on Google. In the academic year 2007/2008 DCU used DORAS for a one-year e-theses pilot with the Schools of Biotechnology and Electronic Engineering. Postgraduates submitting research masters or PhD theses were required to deposit an electronic version of their theses in DORAS. The pilot was successful and from autumn 2008 all research theses submitted for award in DCU must be submitted through DORAS. It uses EPRINTS software, is multidisciplinary and up to 4-2-2010 there were 1454 items on deposit (OPENDOAR, 2010).

NUI Maynooth Eprint Archive http://eprints.nuim.ie/. NUI Maynooth first launched ePrints in 2003 and was the first Irish university to host an institutional repository (Lawton, 2009). It showcases the research output of NUI Maynooth and St. Patrick's College staff and postgraduate students. Its collection has grown steadily from 343 deposits reported in 2006 (NUI Maynooth, 2007) to 2025 up to 21-6-2010 (NUI Maynooth, 2010). ePrints has pre- and post-print research papers made available online by their authors. The papers may be journal articles, conference papers, book chapters, theses, technical reports or any other form of research output (Redmond Maloco, 2007). The resource has excellent statistical reporting capabilities providing anyone who accesses it monthly, daily and even hourly usage statistics, statistics by file type, country and even statistics for top ten keywords and key phrases. The resource has a specialized eTheses function and users can browse by author, department, subject, type or year with advanced and simple searching. NUI Maynooth's ePrints is run by the Library Service with assistance from the Computer Services Department.

Research_online@ucd http://irserver.ucd.ie/dspace/: This is the repository for University College Dublin. Some items are not available as full text. Users can keep track of new additions to the collections by registering for email alerts and RSS feeds. It uses DSPACE software and provides free open access to scholarly research publications such as articles, conferences, theses, unpublished works in the Ecology and Environment, Computers and IT, Business and Economics subject areas. Up to 15-9-2010 Research Online had 1286 items on deposit (OPENDOAR, 2010). Research_Online provides statistics on the global, community, collection and item level.

TARA (Trinity's Access to Research Archive) http://www.tara.tcd.ie/. This site is a university repository providing access to the publication output of Trinity College, Dublin. Some items are not available via Open Access and are only available as metadata (bibliographic record) entries if publisher copyright restrictions apply (OPENDOAR, 2010). Registered users can set up email alerts to notify them of newly added relevant content. The site is well supported with background information and guidance documentation as well as a useful statistics function. TARA uses DSPACE and covers the Science General, Electrical and Electronic Engineering, Geography and Regional

Studies, History and Archaeology, Social Sciences General subject areas and has articles, theses, unpublished works as well as multimedia. As outlined by Brennan (2007) TARA is increasingly being used for preservation purposes to house older publications which have been digitized using high speed scanners by Eneclann, a TCD campus company specializing in high-quality digitisation. These older publications are individually catalogued to TCD library standards by a team of trained metadata inputers with quality control provided by professional cataloguers and subject specialists. This digital archive and its functionality must be seen within the broader context of the development of TARA, which was built with initial funding from TCD's Research Committee to archive and provide open and free access to the scholarly output of Trinity College Dublin, including peer-reviewed papers, electronic theses, and grey literature, and to investigate the archiving of images as primary research materials (Brennan, 2007). TARA has approximately 2512 items on deposit at present.

University of Limerick Institutional Repository: http://ulir.ul.ie/. This relatively new site provides access to the output of the University of Limerick, uses DSPACE and is multidisciplinary. Users may set up RSS feeds to be alerted to new content. Up to 16-4-2010 this repository had 245 items on deposit. It features published and unpublished works of faculty and researchers at the University of Limerick.

Waterford Institute of Technology Repository http://repository.wit.ie/. This is an academic repository for Waterford Institute of Technology (WIT), a third level college in the South of Ireland. It is powered by Eprints 3, free software developed by the University of Southampton. Recent additions can be tracked in Atom or RSS. Not all items are available as open access materials, and some can only be viewed by registered users. The repository has 565 items on deposit (WIT, 2010). This institution has been keen to promote the value of data on repositories and its enrichment by social data through initiatives like OpenID (single sign-on solution to enable a user to authenticate once and gain access to the resources of multiple software systems) and Connotea (a social bookmarking service for academic publications) and their role in the prevention of the information overload that can be a downside of Open Access. (Mulvany & Kane, 2008). They argue that one way of overcoming this overload is to increase the ways in which people may evaluate information by using social data to evaluate existing metadata about an item. As outlined below, social data is about the collection of information about users' preferences that, in aggregate, can be used to provide reliable predictions about what an individual might be interested in.

> For example, in LibraryThing. . . when a user visits a page for a 'work' he or she is instantly able to see what (other) users who bookmarked this work also bookmarked. These items are also likely to be of interest to the visitor who is interested in a specific work. This is because these suggested items represent an aggregation of bookmark data generated by users who all have an interest in this particular item (book). The situation is similar for tagging with uncontrolled keywords/phrases. Users who click on 'fishing' for example are presented with a list of books for which 'fishing' is a prominent tag. (Mulvany & Kane, 2008)

With OpenID the user logs in to a website by authenticating against a 3rd party server and as OPENID is an open standard it can be implemented by anyone. Essentially what WIT are promoting through Connotea and OpenID is the realisation that a scholarly communications system that adopts the features that distinguish many of these social web applications will be far more compelling and useful than a passive search portal and that the adding of a 'social dimension to scholarly communication will have a synergistic effect, increasing the ease with which relevant materials can be sought and discovered by researchers (Mulvany & Kane, 2008).

National and European Initiatives

The benefits of institutional repositories are obvious to institutions. They support research and are a 'shop window' for publications and vital for individual authors as they provide increased impact and as Brennan (2007) has noted, academic and research staff are often measured on their personal research impact, which again is based on citations. This has led Irish libraries to develop a network of federated institutional repositories for Ireland. The Irish Universities Association Librarians' Group has come together under an IUA project to construct this national repository using funding granted from the Strategic Innovation Fund (SIF) of Ireland's Higher Education Authority (HEA). The 3-year project began in 2006 with the aim of making all Irish research output freely and openly available on

the web and of establishing IRs in all Irish universities and to create a national portal that will harvest papers from the IR. The national portal went live in June 2010 and is known as RIAN and accessible at http://www.rian.ie/. RIAN (the Irish word for 'path') houses publications from all of the aforementioned Irish Universities but does not include the institutional repository of Waterford Institute of Technology. It breaks down publications by publication type, funder, institution, year, gives statistics on each publication type and has a sophisticated advanced search where users can tick up to 14 publication types and other boxes for peer reviewed and non peer reviewed content. The content of institutions can be searched simultaneously and RSS is also available.

The Irish Government has identified growth in research as critical to ensuring Ireland's future as a key player on the world stage and essential in underpinning Ireland's investment in its future as a knowledge economy. For individual researchers, studies have shown that as a result of depositing material in institutional repositories citations increase, thus further disseminating the results of research, enhancing the researcher's reputation, supporting future student interest and collaborative partnerships (IUA, 2010). At national level a network of institutional repositories increases the exposure of national research output and allows services, such as enhanced searching and statistics generation to be developed using economies of scale. Services such as RIAN will demonstrate the impact of research to potential funders, who recognise the value of wider research dissemination and the associated research data that supports the reports and conclusions in published research. As Dr. John Hegarty, President of Trinity College Dublin, stated:

> The development of an institutional repository to demonstrate our research activity and achievement is a central element of Trinity's research support infrastructure. …The IReL-open national portal will also play a vital role in communicating the strength of Irish research to a global audience, in attracting the highest quality of research to Ireland and in enabling valuable links with industrial partners. (Irish Universities Association, 2010)

Irish third level institutions are also engaging with Europe on projects such as DART-Europe and DRIVER. DART-Europe is a European organization made up of approximately 120 institutions including some Irish institutions. DART-Europe has created a portal that harvests doctoral theses from IRs. This portal is available at http://www.dart-europe.eu and Hill (2008) reported that approximately 100,000 doctoral theses were harvested into the portal. This rapidly growing portal is not only raising the profile of Irish research in Europe but is also an invaluable resource that can be used by PhD students and researchers.

DRIVER is another European initiative which harvests papers from Irish IRs and is an FP7 funded project with the main objective of building "a virtual, European scale network of existing institutional repositories using technology that will manage the physically distributed repositories as one large scale virtual content source" (IRCSET, 2008 cited in Hill, 2008). Available at http://driver-community.eu it is a multi-phase effort whose vision and primary objective is to establish a cohesive, pan-European infrastructure of digital repositories, offering sophisticated functionality services to both researchers and the general public (DRIVER, 2010). It sets out to build an advanced infrastructure for the future knowledge of the European Research Area and users can search and browse Irish repositories such as DORAS, TARA and others and hence is another route through which researchers can access Irish research.

The NDLR is a national pilot repository system funded by the HEA providing a full ongoing repository service for teaching resources. A key impact of the online NDLR service is to support greater collaboration in developing and sharing of digital teaching resources and associated teaching experience across all subject disciplines and communities of academics and to promote good practice use and re-use of existing resources. By being empowered by the support of communities of academics, staff from different disciplines can share effort and expertise as they raise the bar collectively for how they support their students' learning, embed research in their teaching and potentially embrace partnerships with research and industry, both in Ireland and internationally.

All Irish Universities and Institutes of Technology are partners in the NDLR. A key feature of the NDLR is the nurturing of learning communities and academic communities of practice to encourage successful implementation of the repository. Inter-Institution academic groups work toward improving the subject-specific learning experience and developing online collaborative tools to aid academic communities such as wikis and blogs, video/audio conferencing services. New online collaboration can be suggested and where possible hosted by HEAnet, Ireland's National Education

and Research Network which provides high quality Internet services to Irish Universities, Institutes of Technology and the research and educational community, including all Irish primary and secondary schools. Recent NDLR developments have seen the development of Communities of Practice (CoP's) called 'Smart CoPs'. SMART CoPs are communities in different higher education level disciplines (subject areas), corresponding to the International Standard Classification of Education (ISCED) subject taxonomy, composed of staff interested in the use and application of digital resources and technology in the teaching of their subjects (NDLR, 2010). They are cross-institutional subject discipline related communities associated with locally funded institutional learning innovation projects (LIPS) and cross institutional collaborative projects (LINCs) funded under the auspices of an Oversight Committee. The objective of moving towards project based funding at the local and 'early' collaborative levels is to encourage and sustain the development, sharing and reuse of digital resources, techniques and practices across a range of subject discipline areas once the NDLR moves to become a national service. At present the NDLR has 295 registered users, 23 communities of practice and 3222 items on deposit and so arguably it still has some way to go if it is to have truly national impact and appeal.

NDLR also encourages its users to develop Reusable Learning Resources (RLRs) using such software as Articulate Studio – a set of e-learning software tools which can empower anyone to quickly and easily create engaging courses, presentations, quizzes, and surveys using multimedia content. However the NDLR is not only an online entity and is keen to engage with and encourage communities of practice through running workshops and information sessions around the country on such topics as developing RLRs and Communities of Practice.

Conclusion

It is clear from our review of Irish institutional repositories that collaboration and resources sharing is now a key factor for Irish institutions in making Irish research more visible. We have seen the journey institutional repositories have taken in Ireland from a point where each institution was developing its own repository in the mid 2000s to another more mature and essentially more favorable position where Irish research is now made available on a National Portal and where the collections of each institution can be browsed and searched simultaneously as well as being available through EU repositories. We have also taken a brief look at the NDLR and its use in making teaching resources available online and its role in encouraging academic communities of practice and the development of e learning content. In this new decade Irish universities and institutions must continue to collaborate and share resources and while universities are strongly represented on national portals such as RIAN, Institutes of Technology are not at present. Perhaps this is an area where the Irish Universities Association and the Institutes of Technology could work together to improve the visibility of Irish research. While the NDLR is a highly worthwhile project it is still in development and much promotion and marketing will be required to increase its impact and visibility. What is certain, however, is that with their commitment to collaboration and to the reusability and sharing of resources Irish third level institutions are on the right path.

References

Brennan, N. (2007). Trinity Week Academic Symposium: open access revolution: the local and national context. *Journal of the statistical and social inquiry society of Ireland* vol xxxvi. Retrieved June 2 2010 from http://www.tara.tcd.ie/bitstream/2262/23799/1/very%20latest%2006-07%20version%20brennan.pdf

DRIVER. (2010). *About Driver*. Retrieved June 1, 2010 from http://www.driver-repository.eu/Driver-About/About-DRIVER.html

Hill, R. (2008). *DORAS – increasing the visibility and impact of DCU research.* Paper presented at the International Symposium for Engineering Education 2008, Dublin City University. Retrieved May 25, 2010 from http://doras.dcu.ie/2189/1/hill_rachel_isee08.pdf

Irish Universities Association. (2010). *Open access to university research*. Retrieved June 6, 2010 from http://www.iua.ie/publications/documents/publications/2008/Open_Access.pdf

Lawton, A. (2009). Lenus: from healing God to health repository. *An Leabharlann* 8(2): pp. 26-30. Retrieved June 1, 2010 from http://www.lenus.ie/hse/bitstream/10147/87086/1/Lenus.pdf

Lynch, C.A. (2003). Institutional repositories: essential infrastructure for scholarship in the digital age. *Association of Research Libraries bi-monthly report* 226 (February 2003). Retrieved May 28, 2010 from http://www.arl.org/resources/pubs/br/br226/br226ir.shtml

Mulvany, I. & Kane, D. (2008). *Adding discovery to scholarly search: enhancing institutional repositories with OpenID and Connotea*. Retrieved May 27, 2010 from http://repository.wit.ie/932/1/0R08KaneMulvany.pdf

National Digital Learning Repository. (2009). *Developing Reusable Learning Resources (RLR's) using Articulate Studio '09*. Retrieved May 26, 2010 from ww2.dkit.ie/.../ArticulateWorkshopWorkbook_CB_DKIT_Nov09.pdf

National Digital Learning Repository. (2010). *Statistics*. Retrieved May 25, 2010 from https://dspace.ndlr.ie/jspui/

National University of Ireland Maynooth. (2007). *Annual report of the Librarian* 2005/2006. Retrieved May 22, 2010 from http://library.nuim.ie/about/documents/annrep0506.pdf

National University of Ireland Maynooth. (2010). *NUI Maynooth ePrints and eTheses archive*. Retrieved May 22, 2010 from http://eprints.nuim.ie/information.html

OPENDOAR. (2010). *OPENDOAR directory of open access institutional repositories*. Retrieved June 1, 2010 from http://www.opendoar.org/find.php?search=&clID=&ctID=&rtID=&cID=103&lID=&rSoftWareName=&submit=Search&format=summary&step=20&sort=r.rName&rID=&ctrl=new&p=1

Redmond Maloco, S. (2007). *Maynooths ePrints and eTheses archive*. Retrieved May 22, 2010 from http://eprints.nuim.ie/1023/1/Theology_ePrints.pdf

Waterford Institute of Technology. (2010). *WIT Libraries Repository*. Retrieved May 22, 2010, from http://repository.wit.ie/

Effectiveness Analysis of Electronic Resources at the Hacettepe University

Erdal Coşgun

Hacettepe University, Faculty of Medicine, Department of Biostatistics, 06532 Sıhhiye, Ankara, Turkey. Email: erdal.cosgun@hacettepe.edu.tr

Ayşen Küyük

Hacettepe University Libraries, 06800 Beytepe, Ankara, Turkey. Email: akuyuk@hacettepe.edu.tr

Erdem Karabulut

Hacettepe University, Faculty of Medicine, Department of Biostatistics, 06532 Sıhhiye, Ankara, Turkey. Email: ekarabul@hacettepe.edu.tr

Ergun Karaağaoğlu

Hacettepe University, Faculty of Medicine, Department of Biostatistics, 06532 Sıhhiye, Ankara, Turkey. Email: ekaraaga@hacettepe.edu.tr

Abstract: *It is important to collect and analyze the usage data of electronic databases and periodicals in order to make policies regarding the composition, improvement and more extensive utilization of electronic resources of libraries. The aim of this study is to investigate how efficiently the full text accessible electronic resources of Hacettepe University Libraries are used. For this purpose the usage data obtained from COUNTER Software regarding the electronic databases to which Hacettepe University Libraries subscribed in 2007 were examined. Full-text access to 13,270 journals within 13 electronic databases was found to be 585.843 in 2007. ScienceDirect is the most frequently used database, accounting for 56.1% of all accesses. Some 1,575 journals appear in more than one database and 2,912 journals were never used. A great part of the core journals identified by the Bradford Law are included in the ScienceDirect database.*

Keywords*: Digital libraries; electronic databases; electronic journals; core journals; bibliometrics; Bradford Law; efficiency analysis; Hacettepe University Libraries.*

Introduction

Rapid advances in technology combined with an increase in the number of electronic resources has led to a parallel increase in the number of these sources in library collections. Today, considerable allocations are made from library budgets to electronic resources. With their active role in disseminating information, libraries have been immediately affected by this technological change. Getting acquainted with such new concepts as electronic databases, electronic periodicals and electronic books, libraries have rapidly transformed their collections into electronic resources. The emergence in library collections of electronic resources, which offer the advantages of quick access to information and user-friendliness, has forced libraries to undertake new and different initiatives in collection building and management. This process has made it necessary for libraries to review their collection building policies including selection, provision and evaluation, and to launch utilization and cost analyses to be used in review and decision-making processes.

The recently observed swift increase in the number of periodicals and their prices has led to an increase in the budget share of these publications, which are indispensible, particularly for academic library collections. Hence, libraries are bound to investigate how rational this investment is by assessing to what extent these sources are actually used. Libraries can do this by resorting to such assessment criteria as intensity of collection use, rate at which collections respond to users' needs and identification of most frequently and rarely used materials (Ünal, 2008).

For assessments related to periodicals, there is need to secure documentation and to examine data on in-library use or on articles downloaded from databases. This can form the basis of decisions as to which years of the periodicals should be included or excluded from the collection, to which electronic periodical databases there is need to subscribe or which periodicals in databases should absolutely be included in the collection. Further, among points which should particularly be considered in creating a collection of periodicals, there is the need to identify "core periodicals" which are the *sine qua non* of a collection and to check whether these "must" periodicals change over the years (Ünal, 2008).

The Hacettepe University (HU) Library, at one of the leading universities of Turkey, is aiming to be at the top also in terms of both its resources and the cost effectiveness of these publications. In line with this objective, there have recently been significant achievements in the context of membership to electronic databases providing full texts. The present study assesses, for the pilot year 2007, to what extent electronic databases with full text access are used by the academic staff of HU.

Methodology and Tools

The present study focuses, for the year 2007, on the use of periodicals in full text access databases to which HU has subscriptions. In this context, it was first investigated whether periodicals in databases were effectively used by analyzing the frequency of use. The aim was to expose the effectiveness of databases that provided full text access to articles. For this purpose, the COUNTER software, known as the recording system of databases, was used first. On the basis of this software, the number of uses of full text was found for each periodical. Then, based on these figures, core periodical groups were found for each database and for HU in general. "Core periodical groups" refer to a relatively small number of periodicals that are used frequently. References for publications with the address of HU were identified in the database of Web of Science (WOS).

Bradford Law

Another analysis envisaged under the study was to check whether full article use statistics of databases subscribed to by HU were in conformity with the Bradford Law. The Bradford Law "identifies the dissemination to or distribution of a literature on a specific topic in periodicals" (Garfield, 1980). According to this law, "if scientific periodicals are arranged in order of diminishing returns of articles produced on a specific topic, then periodicals can be divided into a core group of periodicals publishing articles only on this specific topic and multiple groups or zones containing the same number of articles with this core group." (Tonta & Al, 2007) Garfield comments on the Bradford Law as follows: "If you want to compile a bibliography on any topic, there will always be a small core group of periodicals that provide a substantial part (1/3) of articles on that topic or discipline. A second group which is larger than the previous one provides the other 1/3 and finally a much larger third group for the last 1/3." (Al & Tonta, 2007). The distribution of full article use by periodicals was investigated separately for each database and for HU as a whole. The formula suggested by Egghe was used for the Bradford Law. Some values must be known in advance in order to use this formula. These are values p, y_0, r_0 and k, y_0 being the value obtained when the total number of articles in bibliography (A) is divided by the number of selected sub-groups (p) while k is the Bradford multiplier, and y_m stands for the number of articles in the periodical which contains the largest number of articles. Finally, e^Y in the Bradford multiplier is the Euler coefficient = 1,781. (Egghe and Rousseau, 1990).

$y_0 = A / p$

$k = (e^{\gamma} y_m)^{1/p}$

$r_0 + r_0k + r_0k^2 + \ldots + r_0k^{p-1} = T$

from the equation above:

$r_0 = (T(k-1) / (k^p-1))$.

Findings and Discussion

Frequency of Use of Electronic Databases and Resources by Academic Staff

Under the study, data from 13 databases to which Hacettepe University Libraries have subscriptions were obtained (Table 1).

Table 1. Databases for which data available and used in the analysis

Database	Number of journals	%
ACS	37	0.3
Annual Review	41	0.3
ASC	2,436	18.4
Blackwell-Synergy	1,038	7.8
BSC	1,136	8.6
Nature	79	0.6
OVID	288	2.2
Oxford	191	1.4
Proquest	1,694	12.8
ScienceDirect	2,412	18.2
Science Online	1	0.0
Taylor	3,315	25
Wiley	602	4.5
Total	13,270	100.0

Use statistics pertaining to 13,270 periodicals were obtained for these 13 databases. A total of 1,575 of these periodicals appear in more than one database. It should be noted that eight journals that appeared in more than one database were never used at all, even though HU paid at least twice for those journals (e.g., *Acta Archaeologica, Acta Mechanica Solida Sinica, Advanced Powder Technology, Archaeology, Ethnology and Anthropology of Eurasia, Forstwissenschaftliches Centralblatt, Gesunde Pflanzen, Journal of Scheduling, Reproductive Medicine and Biology*). It is observed that the issue of journals included in more than one database has been neglected in studies carried out so far. To solve this problem, the Clementine (SPSS) data mining software was used and these journals and the number of their uses in each database were combined. The frequency of full text article use was 50,06 ± 184,13 for all periodicals.

Of 13,270 periodicals for which data were available in this analysis, 1,575 that appeared in more than one database were excluded and the study was carried out with the remaining 11,695. Out of these 11,695 periodicals, 2,912 (approximately 25%) stood out for zero use, despite their paid subscriptions (sometimes paid for at least twice, see above).

Table 2. Frequency of use of the journals included in the study

# of times full text used	# of journals	%
0	2,912	24.9
1-5	4,047	34.6
6-10	983	8.4
11-25	1,086	9.3
26-50	745	6.4
51-100	638	5.5
101-200	548	4.7
201-500	480	4.1
501-1000	180	1.5
1001 +	76	0.6
Total	11,695	100.0

The number of uses of 13 electronic databases providing full text access to electronic periodicals that the libraries of HU subscribed to in 2007 is given in Table 3. In HU, 585.843 full text articles were downloaded from these databases that year. The majority of these articles (56.1%) were provided by the ScienceDirect database.

Table 3. Distribution of full text use in 2007 at HU of 13 databases (data obtained from COUNTER)

Database	Number of times used	%
ACS	13,275	2.3
Annual Review	3,542	0.6
ASC	12,926	2.2
Blackwell	57,566	9.8
BSC	7,124	1.2
Nature	19,217	3.3
OVID	32,597	5.5
Oxford	22,184	3.8
Proquest	16,664	2.8
ScienceDirect	329,760	56.1
Science Online	3,990	0.7
Taylor	39,059	6.6
Wiley	30,259	5.1
Total	585,843	100.0

Examining the average number of uses of periodicals in electronic databases in the year 2007, we find that the figures for the ACS and Nature databases are higher than others. ASC and BSC have the lowest number of uses per periodical title (Table 4).

Table 4. Use of electronic journals by database

Database	Number of Journals	Number of Full Text Usage					Percentiles		
		Total	Mean	S.Dev.	Min	Max.	25	50	75
ACS	37	13,275	349.3	-	-	-	-	-	-
Annual Review	41	3,542	86.4	101.0	0	490	9.0	64.0	125.5
ASC	2,436	12,926	5.3	11.2	1	284	1.0	2.0	5.0
Blackwell-Synergy	1,038	57,566	55.5	135.5	0	1,377	.0	5.0	44.0
BSC	1,136	7,124	6.27	17.1	1	373	1.0	2.0	5.0
Nature	79	19,217	243.3	389.2	1	2,993	60.0	156.0	300.0
OVID	288	32,597	113.2	209.0	1	2,211	10.0	41.5	135.0
Oxford	191	22,184	116.1	242.7	1	1,386	9.0	22.0	98.0
Proquest	1,694	16,664	9.8	26.8	1	531	1.0	2.0	7.0
ScienceDirect	2,412	329,760	136.7	327.0	0	5,299	1.0	24.5	133.0
Science Online	1	3,990	3,990.0	-	3,990	3,990	3,990.0	3,990.0	3,990.0
Taylor	3,315	39,059	11.8	50.0	0	997	.0	.0	3.0
Wiley	602	30,259	50.3	140.4	0	2,021	.0	6.0	38.0

Core Periodical Groups for Electronic Databases

Analyses were made separately for each database to check conformity with the Bradford Law. It was found that none of 13 databases complied with the law. Considering all periodicals in general, the core periodicals group university-wide for 11,695 periodicals was identified and a test was conducted to check conformity with the Bradford Law. Outcomes are given in Table 5 and Figure 1. In all

databases, there are more periodicals than envisaged by the Bradford Law in Group 1 (the group that constitutes the first 33% of full text use). Further, examining Figure 1 showing the cumulative distribution of total articles belonging to each database by total number of periodicals, we obtain outcomes similar to the ScienceDirect database. In other words, a relatively small number of periodicals out of the total respond to the need for articles.

One possible reason for having no database complying with the Bradford Law could be the fact that the databases comprise periodicals from different areas. It is observed that there is no homogeneity of subject matter. Since database managers strive to include popular periodicals in their series, periodicals from very different areas may come together disproportionately. In general, scientific advances are faster in medical and natural sciences than in social sciences. Considering that most popular periodicals are from the latter, non-compliance with the Bradford Law should be no surprise. While checking conformity with the Bradford Law, the number of groups is accepted as 3 in line with other studies in relevant literature. Nevertheless, conformity was also checked by making 5 and 7 groups, but still no conformity could be found for any database.

Based on this analysis, the first 20 from the core periodical list of HU and numbers for full text usage are given in Table 6.

Table 5. Conformity with Bradford Law for 11,703 journals in the databases to which HU subscribed in 2007

	According to times full text articles used			According to Bradford Law		
	Number of Journals	Number of Papers		Number of Journals	Number of Papers	
Group	n	N	%	n	N	%
1	152	195,259	33.3	16	63,561	10.8
2	539	195,364	33.3	520	290,059	49.5
3	11,004	195,220	33.3	11,159	232,223	39.6
Total	11,695	585,843	100.0	11,695	585,843	99.9
ym = The number of articles in the most commonly used journal		5,559				

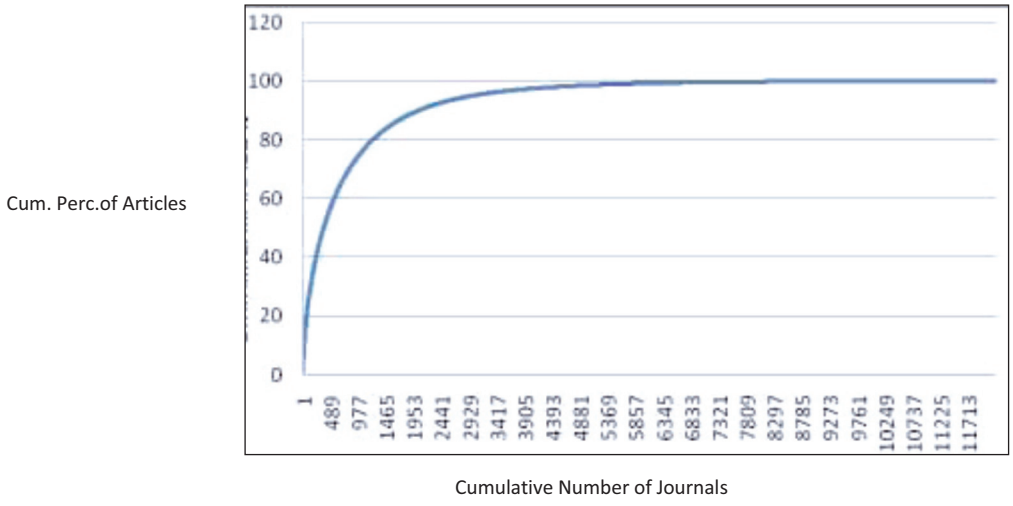

Figure 1. Bradford curve of 11,695 journals in databases to which Hacettepe University Libraries subscribed in 2007

Table 6. Core list of journals for Hacettepe University

Journal Name	Number of Uses
Lancet	5,559
Fertility and Sterility	4,494
Science	3,990
Food Chemistry	3,848
Journal of Chromatography A	3,757
Nature	3,055
Biomaterials	2,907
International Journal of Radiation Oncology*Biology*Physics	2,879
International Journal of Pharmaceutics	2,747
Analytica Chimica Acta	2,347
Journal of the American College of Cardiology	2,320
Journal of Prosthetic Dentistry	2,316
Plastic and Reconstructive Surgery	2,211
Journal of Food Engineering	2,131
Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology	2,027
Journal of Applied Polymer Science	2,021
Journal of Hydrology	1,936
Journal of Pharmaceutical and Biomedical Analysis	1,928
Journal of the American Dietetic Association	1,919
Journal of Chromatography B	1,873

Discussion

HU full text use statistics were evaluated both separately for each of 13 databases and for all databases in general. According to this analysis. the most frequently used database is ScienceDirect with 329,760 articles. The least used is the Annual Review database with 3,542 articles. As for average use per periodical, the ACS database leads the list with 349.34 while the ASC is at the bottom of the list with 5.30. The ScienceDirect database as the most frequently used is in the upper part of the list in terms of average figure per periodical with 136.7. It is clear that this finding can serve as a significant criterion for effective use. It became clear that the equation "the more periodicals = the more uses" may not hold true all the time. The authorities need to conclude contracts with database dealers on the basis of these findings and data.

According to statistics on use, for all 13 databases use is concentrated in a relatively small number of periodicals. These periodicals may therefore be considered as "core" for the databases concerned. Many other studies (Tonta & Ünal, 2007; Bensman, 2005; Tonta & Ünal, 2005; Wiley & Chrzastowski, 2002) have reached the same conclusion. However, the most important point to make here is that it would be a mistake to keep, solely on the basis of use statistics, only frequently used periodicals in library collections and terminate subscriptions to others. Especially in universities teaching in different branches of science, academic staff/student ratio may vary with respect to departments. It is therefore necessary to investigate how important less frequently used and out of core periodicals can be for different departments. The simplest and most effective way of doing this is to administer questionnaires (to assess the effect of periodicals in their respective areas). "Branch of science" oriented questionnaires to be completed by users should be evaluated together with use statistics.

Table 7 gives a comparative account of Bradford analyses in a study conducted in the period 2001-2007 on libraries in Turkey having membership in ULAKBIM for articles downloaded from ScienceDirect and Wiley databases (Tonta & Ünal, 2008a) and our own findings. In neither study do periodical uses comply with the Bradford Law.

According to Table 7, our study points to 69 periodicals in the core group for the year 2007 whereas the other study found this number to be 93 countrywide. However, taken generally, the use behavior of Hacettepe University is still close to the country figure for this particular database. The earlier study did not consider 2007 data for the Wiley database. Still, if comparison is made on the basis of 2006 data (Table 8) the HU use profile is again close to the country profile.

Table 7. Comparison of the group distribution of the journals for the articles downloaded from the ScienceDirect database by member libraries of HU and The Turkish Academic Network and Information Centre (ULAKBIM)

Group	Our Study 2007	Tonta & Ünal 2008a
1	69	93
2	215	257
3	2,128	1,663
Total	2,412	2,013

Table 8. Comparison of the group distribution of the journals for the articles downloaded from Wiley database by member libraries of HU and ULAKBIM in Turkey

Group	Our Study 2007	Tonta & Ünal 2008a
1	14	19
2	44	57
3	544	361
Total	602	437

Another study conducted in Turkey (Tonta & Ünal, 2007) found that researchers in universities having membership in ULAKBIM provided for a large part of their need of articles by referring to relatively few core periodicals available in electronic periodical databases in ScienceDirect, SpringerLink, WileyInterScience and EBSCOHost of ULAKBIM. Our study, on the other hand, found that the core periodical group of HU comprised 160 periodicals. Examining these periodicals by their fields we find that they are mostly concentrated in medicine and medical sciences. Hence, considering that the list does not encompass all branches of science in the university, it would be incorrect to treat the list as the core periodicals group of the university.

Conclusion and Suggestions

Creation and maintenance of library collections is difficult and requires costly investment. Hence for their effective utilization there is need for keeping correct use statistics and analyzing these statistics properly. Otherwise it may happen that unsound investment is made in periodicals which are not used at all or used only rarely or double payment is made for the same periodical since a single periodical can be marketed to libraries by more than one database. At the stage of determining frequencies of use, there is a need to take into account some very important points. Since databases are accessible from different web addresses, a single periodical can be counted more than once as if different periodicals are in question, which has a misleading effect on frequency of use as well as showing the number of periodicals to be greater than the actual number.

When the title of a specific periodical appearing in different databases starts with an upper case or lower case letter depending on the choice of the database, or if a title is with or without suffixes in English, the list of periodicals obtained from the COUNTER software may turn out to be longer than it actually is. For example, the periodical cited as *Lancet* in one database may be cited as *The Lancet* in another, or *Kidney International* in one database becomes *Kidneyinternational* in another, which may lead to double counting. These factors should be analyzed carefully. This analysis should be conducted not by hand but through methods like text mining which are technically more sound and which minimize the possibility of error. For efficient utilization of such a costly investment, researchers who will utilize relevant services have to fulfill their responsibilities just as library management rendering relevant services has their specific duties to perform. It is a must for researchers to learn more about this immense source of information that they can access in seconds by using their computers, to be determined and willing to use accessible electronic databases more correctly and effectively and to receive training if necessary.

Given that 1,575 of 13,270 periodicals which are found in HU Libraries and are covered by this study appear in more than one database; that 2.,912 periodicals (25% of total) have never been used; and that 8 of these zero-use periodicals are marketed by more than one database, it appears that it would be economically more sound to subscribe to periodicals individually rather than on the basis of

databases. This suggestion is further supported by the unequal distribution of frequency of use of periodicals in a database and the rather pronounced nature of this inequality.

Meanwhile, the concentration of the core group of periodicals identified in the study in a few disciplines indicates that statistics of use alone would not be sufficient in a study of this kind. In efforts to create a core group of periodicals for an academic institution like HU, there is a need to take into account such other criteria as branches of science in the institution and number of users in respective branches. In other words, a core group of periodicals should be established for each and every branch of science. It is considered that the findings of the present study on HU, one of the leading universities of Turkey, are quite important for the relevant literature. This importance derives from the fact that the university encompasses a wide range of academic units including Medicine and Health Sciences, Science and Social Sciences, Engineering, Law, Fine Arts, Music and Stage Arts; that sources of information provided by the library are quite rich and diverse; and that the group of users is quantitatively large. Conclusions drawn from the model of such a large and diversified university can serve as a model for many other academic institutions.

Acknowledgements

Funding: Supported by The Scientific and Technological Research Council of Turkey (Project Number: 108K350). We also thank Dilek Şahin, Eda Köse and Pınar Al for their valuable contributions.

References

Al, U. & Tonta, Y. (2007). Tam metin makale kullanım verilerinin bibliyometrik analizi. In S. Kurbanoğlu, Y. Tonta & U. Al (Yay. haz.). *Değişen Dünyada Bilgi Yönetimi Sempozyumu. 24-26 Ekim 2007. Ankara. Bildiriler* içinde (pp. 209-217). Ankara: Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü.

Bensman, S.J. (2005). Urquhart and probability: the transition from librarianship to library and information science. *Journal of the American Society for Information and Technology,* 56(2), 189-214.

Egghe, L. & Rousseau, R. (1990). *Introduction to informetrics: quantitative methods in library, documentation and information science*. Elsevier, Amsterdam.

Garfield, E. (1980). Bradford's Law and related statistic patterns. *Current Contents*, 19. 5-12.

Tonta, Y. & Al, U. (2007).Cumhuriyet Dönemi Makaleler Bibliyografyası 1923-1999: Eleştirel Bir Değerlendirme. In S. Kurbanoğlu. Y. Tonta & U. Al (Eds.). *Değişen Dünyada Bilgi Yönetimi Sempozyumu. 24-26 Ekim 2007. Ankara. Bildiriler* içinde (pp. 105-112). Ankara: Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü.

Tonta, Y. & Al, U. (2008). Türkçe Makalelerin Dergilere Dağılımı ve Bradford Yasası. *Bilgi Dünyası*, 9(1), 41-66.

Tonta, Y. & Ünal, Y. (2005). Scatter of journals and literature obsolescence reflected in document delivery requests. *Journal of the American Society for Information Science and Technology*, 56(1). 84-94.

Tonta, Y. & Ünal, Y. (2006). The impact of electronic journals on document delivery services. B. Martens & M. Dobreva (Eds.). *Digital Spectrum: Integrating Technology and Culture Proceedings of the 10th International Conference on Electronic Publishing, Bansko. Bulgaria, June 14-16, 2006* (pp 329-338). Sofia: FOI-Commerce.

Tonta, Y. & Ünal, Y. (2007). Dergi kullanım verilerinin bibliyometrik analizi ve koleksiyon yönetiminde kullanımı. In S. Kurbanoğlu, Y. Tonta ve U. Al (Eds.). *Değişen Dünyada Bilgi Yönetimi Sempozyumu. 24-26 Ekim 2007, Ankara. Bildiriler* içinde (pp. 193-200). Ankara: Hacettepe Üniv. Bilgi ve Belge Yönetimi Bölümü.

Tonta, Y. & Ünal, Y. (2008a). *Dergi kullanım verilerinin bibliyometrik analizi ve koleksiyon yönetiminde kullanımı*. TÜBİTAK SOBAG 106K068 nolu proje raporu.

Tonta, Y. & Ünal, Y. (2008b). Consortial use of electronic journals in Turkish universities. In L. Chan & S. Mornati (Eds.). *Open Scholarship: Authority, Community and Sustainability in the Age of Web 2.0: Proceedings of the 12th International Conference on Electronic Publishing, Toronto, June 25-27, 2008* (pp. 203-216). Toronto, Canada: International Conference on Electronic Publishing (ELPUB).

Ünal, Y. (2008). *Dergi kullanım verilerinin bibliyometrik analizi ve koleksiyon yönetiminde kullanımı*. (Unpublished PhD dissertation). Hacettepe Üniversitesi. Ankara.

Wiley, L. & Chrzastowski, T. (2002). Interlibrary Loan Assessment Project II: revisiting statewide article sharing and assessing the impact of electronic full-text journals. *Library Collections and Technical Services*. 26,19-33.

What Makes a Great [Mobile | Digital] Art Exhibition?

Mícheál Mac an Airchinnigh

School of Computer Science and Statistics, University of Dublin, Trinity College, Dublin, Republic of Ireland. Email: mmaa@cs.tcd.ie

Glenn Strong

School of Computer Science and Statistics, University of Dublin, Trinity College, Dublin, Republic of Ireland. Email: Glenn.Strong@cs.tcd.ie

Catherine Giltrap

Curator of the Trinity College Art Collections, University of Dublin, Trinity College, Dublin, Republic of Ireland. Email: catherine.giltrap@tcd.ie

Shane Brennan

School of Computer Science and Statistics, University of Dublin, Trinity College, Dublin, Republic of Ireland. Email: brennash@scss.tcd.ie

Mike Brady

School of Computer Science and Statistics, University of Dublin, Trinity College, Dublin, Republic of Ireland. Email: brady@cs.tcd.ie

Abstract: *Passive reception and consumption of art is a given, in our times. Artists produce. Spectators consume. At the nexus stands the curator who chooses the product and the exhibitor who provides the space for consumers. This natural hierarchy also tends to colonize the digital space. But, in the digital world, much of the functioning of the hierarchy has become democratised. The meeting place of exhibited art moved from the physical to the virtual online. Not everyone can visit, say, Istanbul Modern museum. It ought to be possible in principle for everyone to be able to visit "Istanbul Modern Digital" museum. The next stage of digital democracy, already upon us since early 2010, is the mobile art lover, mobile in the sense of being free from being tied down in one place and being able to choose what to see, where to be, and when to do it: early morning, late at night; in the plane, on the train, in bed, in class.*

Learning is for everyone. It is what makes us human, to continue to learn. Learning takes place best when one is active. In the context of the Mobile Digital Art Exhibition, we have explored ways in which to enhance the experience of the curator as "everyman" and everyman has potentially the opportunity to construct a mobile digital art exhibition, even one such as the "Museum of Innocence" in the manner as described by Orhan Pamuk. Our hero in this story is the self-curator.

Keywords: *Flickr Gallery; GoGo; Google Goggles; mobile devices; QR code tags; self-curator.*

Introduction

In the modern world of the digital cultural heritage, it is to be taken for granted that one will want/need to organize an exhibition, to show off some particular collection of art works. There are three principal types of such exhibitions. The first type is organized along classical lines. The art works are arranged in a physical space. Often such a space is a building, an art gallery for instance. The Trinity College Dublin (TCD) Art Collection was recently exhibited in such a manner. But it is also possible to arrange an art exhibition outdoors, in an unconfined physical space. Thirdly, in our times, one can facilitate a mobile digital art exhibition ***(MDAE)***, accessible to any interested party anywhere in the world. A fundamental research question for the (self-)curator of such a mobile digital

art exhibition is simply this: How can I prototype my ideas and get feedback? A second fundamental research question is at the heart of this paper: How can I facilitate and augment self-curation in the mobile digital world? For example, one might imagine a graduate/staff member of TCD who participated in the TCD College Art Hire Scheme (Trinity College Dublin, 2010), wishing to self-curate those works of art which formed the backdrop of her/his academic life time in the College. This, of course, can only be done via the (mobile) digital art exhibition, presuming, of course, that the art works are digitised.

One can see how this fundamental idea extends to all those who ever visited an art gallery, bought exhibition catalogues, marked mentally/physically those works to which they were attracted and who would dearly love to have the facility to look over them again. Such a person potentially belongs to that category of the digital art self-curator. The technology now exists to make such (mobile) digital art self-curation a real possibility. *En passant*, it is noteworthy that the recent magnificent Gauguin Exhibition – "Gauguin: Maker of Myth", in the Tate Gallery London, turned out to be, for some, a complete disaster, leading to a neologism: "Gallery Rage" (Thorpe, 2011). One could also imagine a visitor to the Gauguin Exhibition having a personal mobile digital art exhibition selection with her/him and an indication of current crowd size around any specific piece at any given place and time. This might facilitate access to the artwork and avoid what has been termed "art kettling."

Our current research follows on directly from the work done in an earlier paper concerning four art galleries presented under the theme of "Social networks and the national art gallery (Dublin | … | Sofia)" (Mac an Airchinnigh & Strong, 2010). The galleries, in order, were "The National Gallery of Ireland" [**NGI**] (Dublin) ("National Gallery of Ireland," 2011), "The National Art Gallery" [**NAG**] (Sofia, Bulgaria) ("National Art Gallery Sofia," 2008), "Museo Nacional del Prado" (Madrid, Spain) ("Museo Nacional del Prado," 2010), and "The Ateneum Museum" (Helsinki, Finland) ("Ateneum Art Museum," 2010). Of these, the NGI did not afford any opportunity for the folk on the WWW to see the full range of the digital images of the art collection (April 2010). Now, one year on, the NGI compares well with state of the art digital facilities worldwide. However, in the context of a ***mobile digital art exhibition (MDAE)***, there is much yet to be done, and we will *assume* that any remarks made, may apply, *inter alia*, to the Ateneum and El Prado. Our discussion on MDAE will apply to all four galleries, precisely because of the nature of "state of the art" digital art galleries in 2011. On the other hand, there seems to have been one major change in the digital status of the NAG in Sofia. Hitherto, the small exhibition of the images of paintings online had been served by both Bulgarian and English pages. At the time of going to press the English pages are no longer accessible. We will make some suggestions on how one might make the leap from the NAG status quo to some sort of rudimentary mobile digital art exhibition. Paradigmatically, this will apply to all others.

Another fundamental issue is the determination of the nature of a mobile device. That said device is mobile, i.e., it can move around (whether autonomous or not), is a given. But, in the context of this paper, the mobile device must be able to connect to the internet wherever it finds itself, whether on the move or at rest. For example, any laptop, such as a MacBook Pro, is by definition such a device. On the move, say on a train, the aforementioned device can connect to the internet, if *tethered* with an iPhone. The concept of tethering means that the mobile phone (such as the iPhone) provides the internet connection (3G) to the laptop (such as the MacBook Pro). The iPhone and iPad are, by definition, such devices. Such interconnecting of multiple mobile devices determines the platform for total stability of mobile dynamically digital self-curated (and otherwise) art exhibitions. We will have more to say on this matter of a mobile device later in the paper.

Curator, Self-Curation

One of the significant prototyping tools for the would-be mobile digital art self-curator is the Flickr Gallery (Flickr, 2011). The Flickr developers "specifically limited the number of photos … that can be added to each gallery to 18 … [in order] to give our members an opportunity to engage in activity that is similar to what a curator of a gallery or museum might undertake." To illustrate this significant and easy to use Gallery feature we have constructed two, one with the title "ElPub2011 Mobile Digital Art Exhibition" (Orela, 2011a) and the other with the title "ElPub2011 National Gallery of Ireland" (Orela, 2011b). The mechanics are simple: "To create a Gallery just go to a photo or video that you want to add and click the "Add to a gallery" link within the Actions menu. From here you can create a new gallery or add the image to an existing one." However, the would-be self-curator

clearly needs some idea, some theme, even to begin. There are millions of photos on Flickr. The first author has a Flickr account and is administrator for the "ЖЕНА (woman) pool". At the time of writing, there are 1153 members and 14,322 photos. To construct a well-themed "Gallery Exhibition of 18 images" from such a pool is a daunting task. One needs to have some significant guidelines on how to choose those 18 images. A good starting point is the essay "Show and Tell" by Robert Storr, artist and critic who serves as dean of the School of Fine Arts at Yale University (Marincola, 2006, p. 172).

Group Exhibition

"Selection is the initial, and, in many ways, the touchiest, stage…" (Storr, 2006, p. 21).

Shall we pick the photos from one individual Flickr-ite or shall we pick one photo from 18 Flickr-ites in the group according to a specific theme or idea? How shall we find the 18 photos among the 14,322? A typical gallery might feature "woman as dancer" (Orela, 2011c). Clearly, if the photos are tagged appropriately with "dancer", ("танцьорка" in Bulgarian), then we have a much smaller pool to choose from. Currently, on Flickr, there are 92 photos tagged with танцьорка and 1,177,382 photos tagged with "dancer". One does not presume that English is to be the sole linguistic access mechanism. "Google translate" does a reasonably good translation for a large variety of languages and in the first stage of our research, translation of typical Flickr tags is more than adequate. However there are other technical issues still to be resolved. Currently, at the time of writing, the Flickr app for the iPhone, does **not** recognize Bulgarian tags such as "танцьорка." Nor does there seem to be a way in which to view a Flickr gallery on the iPhone. One presumes that this is also currently the case with the iPad. It is clear that the first wave of certain apps for mobile devices are crude from the point of view of universal (language/feature) access in comparison with the apps for laptops. On the other hand, Google Goggles (GoGo, a useful needed neologism) is now built into the Google search app for the iPhone (and presumably for all other mobile platforms). One takes a photo of an image such as might be found in an Art (Exhibition) Catalogue or online, say Flickr. GoGo then analyses the image and returns the result of an internet search. Either an exact match will be found or, more interestingly for our purposes, other images which match certain criteria. For example, using GoGo to photograph the image of "Virginie" (1883) by Albert Edelfelt (Edelfelt, 1883; Mac an Airchinnigh & Strong, 2010) from a catalogue will return online versions and other related material. But we can do much better with the GoGo technology.

Figure 1. An iPad layout for a mobile digital art exhibition (Mike Brady)

Showing is Telling

"The primary means for 'explaining' an artist's work is to let it reveal itself. . . Space is the medium in which ideas are visually phrased" (Storr, 2006, p. 23). This forthright principle has been put into practice by devising a simple online exhibition (hosted on Flickr) whereby only the image is shown (Orela, 2011d). All other telling aspects are only accessible by means of a **Q**uick **R**esponse (QR) code tag (Mac an Airchinnigh & Strong, 2010). The original precursor for such a QR tagged exhibition was a 3-person MSc Research Project, with the title "Smartphones Let Loose—The Museological Potential of Wireless Technology" (Aggarwal, Molines, & Wei, 2010). For a variety of reasons, the experiment/exhibition had to be conducted in the Lloyd Building of Trinity College Dublin. Full high-resolution prints of a selected number of paintings from the National Gallery of Ireland were bought, mounted behind glass, and made visible to passers-by (students, academic-and-other staff, visitors). Each print was tagged by a shortened URL (bit.ly) and also by the corresponding QR code of that URL. At the time (in 2010), GoGo was in its infancy. One year on, we note that there is no longer any need for the URL. The QR code tag is fully sufficient.

One of the major concerns that we still have is in the general field of aesthetics. This concern may be described in connection with the document shown in Fig. 1 below. Does not the juxtaposition of a QR tag "interact" with the image to which it relates? Does not this interaction "detract" from the image? Is there an appropriate location for the QR tag with respect to the kind and dimensions of each image?

Mobile Devices and Art Apps

Apple's iPad (Apple**,** 2011) (which we imagined might be called iTablet (Mac an Airchinnigh & Strong, 2010) in early January 2010) paved the way for the newly realized generation of tablet devices. In the context of the Trinity College Art Collection Exbibition, discussed below, an iPad Art Application — the TCD ART APP — was designed with a main image pane to show the specific art work and side panel to the left giving a variety of information on the piece, for example, location, provenance, and so on. The iPad application is a general purpose media-presentation 'app' that can display high-quality images, text, video and audio. Briefly, the app presents the user with a selection of images of a fixed size. The images may comprise or contain video and audio sequences, as well as high quality graphic and textual elements. Media formats include H.264 movies (Wikipedia, 2011a), AAC audio (Wikipedia, 2011b), JPEG 2000 (Wikipedia, 2011c), GIF and TIFF images and PDF. In addition, pages can be laid out using standard HTML and CSS. It is also worth noting that there were issues with the performance of the JPEG 2000 interaction with PDF documents on iPad 1 ("Apple Discussions, Topic: iBooks on iPad not showing images in PDFs," June 26, 2010). One presumes the issue is resolved now on both iPad 1 and the newer iPad 2.

Compared with a web-based approach, the app offers the following advantages:

- All media—images, videos, audio, text, graphics—are contained in the app. The user has instant access to any of the media in the app, without being at the mercy of a network connection. Artists are free to concentrate on providing the highest quality media, unconstrained by network speed or capacity considerations.
- The app presents a fixed format for the presentation of media.

For distribution, iPad apps must be submitted to the iTunes App Store, and must then be approved by Apple Corporation. It is understood that apps are vetted for compliance with technical guidelines and for general suitability and utility. They also require an audience suitability rating. Generally speaking, the time from submission to approval is less than two weeks. The submission process is fully automated from the developer's point of view.

Playbook is a very nice title for the RIM tablet (Halliday, 2010). Such a tablet is a general portable flat *universal* computing platform that generalizes the so-called e-book readers such as the Kindle (Driver and Vehicle Licensing Northern Ireland, 2006). A different case to the 'interactive gallery guide' previously imagined. Instead of using a camera to capture images and deliver descriptions (narratives, questions, whatever) we imagine an instructor-driven (classroom) environment. The larger surface of the tablet allows a smoother interaction with the documents, with greater ease of manipulation for small components. Using a network-based document share, the

instructors can publish individual documents (images, text, etc.) so that they appear on the students' tablets as needed, and a student's interaction with the document can be captured (either for student assessment or to guide the teaching process). Rich documents might be authored using the upcoming HTML5 standard and synchronized using a cloud-store model.

In the remainder of the paper we give four examples of self-curation and exhibit some of the problems encountered. The activity of self-curation was fun, from a technical point of view.

Case Study 1: The Trinity College Art Collection


QR tag 2


QR tag 1

In December 2010, the Curator, Catherine Giltrap, published a book, entitled "George Dawson: An Unbiased Eye" (Giltrap, 2010a) to celebrate 50 years (since 1959) of activities undertaken by George Dawson, the Professor of Genetics at Trinity College Dublin. A succinct account of this undertaking appeared in The Irish Times newspaper in October 2010 (Dunne, 2010). However, in conjunction with this classical activity, there was a special effort made to engage with the digital world at the same time. Three specific activities are to be noted in this respect: (1) the development of an App for the iPhone and iPad to illustrate some key art works from the collection; (2) the organization of a walkabout tour of art works whereby QR codes were used as digital links to information on the College Web about the art work in question (See Fig.1 below); (3) the hijacking of the classical Newgrange Winter Solstice (2011) spectacle by overlaying the event with a Digital Solstice of Art taken from the collection. The latter has been postponed until December 2011 (Giltrap, 2010b).


QR tag 3

To illustrate the "showing is telling" principle, a one-page art pamphlet was created (Fig. 2). There are four pieces of art shown, centred in each row: 1. *Le Diable* (Appel, 1956); 2. *Oiseau dans le Feuillage* (Braques, 1961); 3. *Study '66* (Farrell, 1966) and 4. *Kite!* (Scott, 1974). There is a short text giving some standard information on the artist, the year, the materials, the provenance and so on. There is a QR code tag associated with the piece of art. Anyone with a smartphone and QR reader will be able to access the full text associated with the artwork by taking a photo. Using the GoGo feature on the iPhone, each QR code tag (eventually) led to the appropriate web page in Trinity College Dublin (2011-01-04). Interestingly, GoGo produced alternative possibilities to the interpretation of the QR code tag. Nothing is perfect!

The display pattern of the alternating QR code tags on the art pamphlet has an aesthetic appeal. But, behind this façade there is a more serious purpose: 1. The minimum printed size is approximately a square of side 1 cm, and 2. There must not be another QR code tag within "half a tag width" (Brennan, 2010). Now that one can see the way in which to present art/images with QR code tags, one will be interested in experimenting for oneself. The basic (missing) link is, of course, the QR code generator. A typical search provides multiple possibilities. One choice leads to the Kaywa QR code generator (Kaywa, 2011). Any QR reader will (ought to) return the URL and web page of "2. Oiseau dans le Feuillage" above (Braques, 1961a, 1961b). On the other hand, one might start with the print of "Oiseau dans le Feuillage" as shown in "George Dawson: An Unbiased Eye", Fig.20, p.38 (Giltrap, 2010a). Application of GoGo returns the primary result as a web page of the Stadtmagazin Echo Münster (Stadtmagazin Echo Münster, 2011), a result which might seem to be surprising for the novice.

To the best of our knowledge, this was the first public experiment of the use of the smartphone to access more complete information on the artworks in an exhibition (in Ireland). Naturally, similar such undertakings might be expected elsewhere (Chu, 2010). The design was not perfect in the sense that there was readable formal identification printed alongside both small size (full-colour) images of the artwork and the QR tag. Specifically, given the identification of the art work there was no real incentive (other than technological curiosity) to use GoGo.

George Dawson: An Unbiased Eye. Modern and contemporary art at Trinity College Dublin since 1959

19 November – 19 December 2010, The Royal Hibernian Academy, Ely Place, Dublin 2

SMARTPHONE INTERACTIVE EXHIBITION TRAIL - preview

Instructions: Smartphone users can download a free QR tag reader app to scan the following. Simply open the QR app once downloaded. This automatically opens the smartphone camera. Hover the camera over the black and white tags. Your phone will be directed to a page on the Trinity College Dublin Art Collections website where you can access more information on the artist and the work. Enjoy!

Direct access to the web pages is also available at www.tcd.ie/artcollections/exhibitions/index.php

Special thanks to Prof. Micheál Mac an Airchinnigh and Shane Brennan, School of Computer Science, TCD, for their friendly assistance and support – both creative and technical

Figure 2. QR code tags for smartphones

Case Study 2: The National Art Gallery (Sofia, Bulgaria)

QR tag 4

One might imagine the use of a smartphone camera on an original art image in a gallery. For obvious reasons (one of which is that of copyright), many galleries would not facilitate such a thing. But, even if allowed, there might be artworks which would resist recognition (at least in the short term). For example, using GoGo directly on the digital image "*Composition – Female Figure III*" circa 1950, by Nicolai Abracheff (Abracheff, 1950), does not result in an appropriate match. One hypothesis for the failure might be that the image is formally too abstract to be easily recognizable (as that of a woman in red).

There is a very simple and effective solution to this problem, now to be explained in some detail in the practical context of the recent Trinity College Art Collection exhibition, outlined above. One generates a QR code tag ("2D Code", 2011) for the URL of the web page on which the digital image of the painting is located. Then the image and the QR code tag are collocated, such as on Flickr (Mihal Orela (Orela, 2011a). Finally, use of GoGo will bring one to the web page in the National Art Gallery (Sofia). The results of this practical research are electronically published in the form of a Flickr Gallery entitled "ElPub2011 Mobile Digital Art Exhibition" (Orela, 2011a).

We can also turn this mobile digital art collection into a simple game: "Spot the odd one out." Specifically, one of the 18 images chosen for the Flickr Gallery does **not** belong to the collection of images on the National Art Gallery (Sofia) web site. The goal is to determine which one and then to determine where it is located. It is worth noting, *en passant*, that the number of images per web page of the National Art Gallery (Sofia) is 12 and 12 = 4x3 is a nice number (months of the year, hours of the day,…). The image paning of a tablet (iPad, iPhone,…) to facilitate a 4x3 or (3x4) exhibition has great aesthetic appeal.

Table 1. Construction of QR tagged art works

Table	URL	QR tag	Image	Image Tagged
NAG Sofia	http://www.nationalartgallerybg.org/index.php?l=55&id=21 [2011-03-31]	QR tag 5		
NGI Dublin	http://onlinecollection.nationalgallery.ie/media/view/Objects/6379/14653?t:state:flow=314462a0-8cf2-4e79-88c1-a50ec9de0b57 [2011-03-31]	QR tag 6		
Museum of Innocence	http://www.flickr.com/photos/mihalorel/5363085749/ [2011-03-31]	QR tag 7		

Case Study 3: The National Gallery of Ireland (Dublin, Ireland)

At the ElPub 2010 Conference in Helsinki we reported on the then status of the digitized collection of the National Gallery of Ireland ("National Gallery of Ireland," 2010). Now one year later ("National Gallery of Ireland," 2011) it is gratifying to see that it compares well with other world class "National" Art Galleries. One feature of particular relevance is the ability to create online one's own collection of images under the rubric of "My Collection" (National Gallery of Ireland My Collections, 2011).

Let us consider the art work entitled "The Meeting on the Turret Stairs, 1864" by Frederick William Burton (Burton, 1864), a popular piece. Use of GoGo currently gives a good variety of URL choices, and that which takes one to the National Gallery of Ireland website is currently on the 5th page of results. . . Undoubtedly over time, as people get used to exploring the NGI web site and using its "My Collection" feature, the NGI URL will rise further. But in the context of a Mobile Digital Art Exhibition, we can greatly augment the standing of the NGI. There is a large choice of images. It is possible to create an external Flickr Gallery (Orela, 2011b) where each image from the NGI collection is given a suitable precise QR tag. An example is shown in Table 1 above.

Case Study 4: The Digital Museum of Innocence (Everywhere for Everyman)

Wanted! Curator for the Museum of Innocence! The most recent news concerning the "real" Museum of Innocence is that it will be in place by Spring 2011 ("Where is the Museum of Innocence?," 2011). One will see. Orhan Pamuk's Museum of Innocence is founded upon his earlier work, "Istanbul, Memories of a City," (Pamuk, 2005). His "prolonged study of these photographs [each one assigned its place in the museum] led [him] to appreciate the importance of preserving certain moments for posterity" (Pamuk, 2005, p. 13). And the Istanbul book is full of photographs. It is also full of his own personal memories up until the age of 18 when he abandoned the career of artist to undertake the career of writer. Taking the two books together one sees how memories accumulate in the things preserved. For our final mobile digital art exhibition we will construct a simple "game"/"play" based on a Flickr Gallery of 18 images. The theme will be the "Seven ages of man": infancy, boyhood, lover, soldier, justice, pantaloon, second childishness (Wikipedia, 2011d, 2011e). **Rule 1**: One will universalize the "ages of man" to apply equally to woman in our times. **Rule 2**: One of the ages will be dropped. It matters not which one. **Rule 3**: There will be three images for each of the six ages. **Rule 4**: Each image will carry the same Creative Commons license. **Rule 5**: Each image will carry the appropriate QR tag. **Rule 6**: Each image *may* carry the designated age of (wo)man in an appropriate aesthetic form. Such game/play may initially be considered to be similar to a game of solitaire. However, with a little more thought and corresponding rules and moderation, one can deduce that the game might readily evolve into a "distributed" family game played out digitally over the web.

Conclusions and Suggestions

Disappearing digital data: In the short space of time between the initial submission of the paper in late January 2011 and delivery of this, the reconsidered final form, to be published in early April 2011, that is to say in the space of little more than two months, some electronic publication (elpub) data have been invalidated. Experience of the digital world, since at least the millennium year 2000, has shown that digital data (web pages, for example) can and do disappear from public access. In our case, in the development of the Flickr Gallery entitled "ElPub2011 Mobile Digital Art Exhibition," we used the English form of the URLs to locate the images of the artworks in the National Art Gallery, Sofia and generated the corresponding QR tag encodings. We certainly did **not** expect that these Gallery pages would disappear, that is to say become inaccessible. Consequently, any QR tag encoding of such page addresses would become invalid. Therefore, use of Google Goggles on such QR code tags would not lead to the desired destination. For example, in exposing raw URL data in Table 1 above we discovered the need to replace the English page:
http://www.nationalartgallerybg.org/index.php?l=60&id=21
with the corresponding Bulgarian page:
http://www.nationalartgallerybg.org/index.php?l=55&id=21 [2011-03-31]

Contextuality of QR tags: For some, the QR tag, being machine oriented, and therefore intrinsically alien to humans, is ugly. Placed in proximity to an artwork, the QR tag has a jarring effect. One cannot be sure that the decoding will be "suitable" until the QR tag is decoded.

QR tag 1: still active ("Curator's Role, College Art Collections: Trinity College Dublin,") 2011-04-03
QR tag 2: still active ("Georges Braque 'Oiseau dans le feuillage' (1961) lithograph on newspaper edition 15 of 50, George Dawson: An Unbiased Eye,") 2011-04-03
QR tag 3: disappeared early 2011 (Stefan Ivanov, 1875-1951); Bulgarian version exists (Иванов, 1875-1951) 2011-04-03
QR tag 4: still active (Абрашев, 1897 – 1962) 2011-04-03
QR tag 5: is exactly the same as QR tag 3, 2011-04-03
QR tag 6: still active (Frederick William Burton, 1864) 2011-04-03
QR tag 7: still active 2011-04-03

Acknowledgements

We are very grateful for the insightful remarks of the referees. The first remarked specifically on the need to elaborate further on the context of the concept of the digital rediscovery of culture. The second showed us that we needed to emphasize the user-driven desire to curate one's own mobile collection. To make this focus clearer we introduced the term self-curate to designate the fact that it is the user who is at the centre, not the institution and not the seller/provider of fixed-content art apps. Happily, both needs are mutually supporting.

References

2D Code (2011). Retrieved Jan 9, 2011, from http://www.the2dcode.com/qr-code-generator

Abracheff, N. (1950) . Composition – Female Figure III, circa 1950. Retrieved Jan 9, 2011, from http://www.nationalartgallerybg.org/index.php?l=60&id=3

Aggarwal, A., Molines, S., & Wei, L. (2010). Smartphones Let Loose - The Museological Potential of Wireless Technology. from aggarwal.arnav@gmail.com, clafou@gmail.com, weilai.irl@gmail.com

Appel, K. (1956). *Le Diable*. from http://www.tcd.ie/artcollections/exhibitions/appel.php last access 2011-01-04

Apple. (2011). Apple iPad. 18-by-24cm.

Apple Discussions, Topic: iBooks on iPad not showing images in PDFs. (June 26, 2010). Retrieved April 3, 2011, from http://discussions.info.apple.com/message.jspa?messageID=11875744

Ateneum Art Museum. (2010). Retrieved January 20, 2010, from http://www.ateneum.fi/

Braques, G. (1961). *Oiseau dans le Feuillage* from http://www.tcd.ie/artcollections/exhibitions/braque.php last access 2011-01-04

Brennan, S. (2010). Minutes of Meeting - 2/11/2010. School of Computer Science and Statistics, Trinity College Dublin.

Burton, F.W. (1864). *Hellelil and Hildebrand, the Meeting on the Turret Stairs*. Retrieved Jan 16, 2011, from http://www.nationalgallery.ie/en/aboutus/Rights and Reproductions/Print Sales/Emotive Scenes/Turret Stairs.aspx

Chu, T. (2010). Retrieved April 2011, from www.core77.com/blog/news/the_opening_of_design_museum_bostons_creative_capital_exhibit_18124.asp

Driver and Vehicle Licensing Northern Ireland. (2006). Retrieved June 19, 2006, from http://www.dvlni.gov.uk/

Dunne, A. (2010). Anchorman, Pop Tarts, Picasso . . . *The Irish Times*. October 7, 2010. Retrieved May 29, 2011 from http://www.irishtimes.com/newspaper/features/2010/1007/1224280562204.html

Edelfelt, A. (1883). Virginie. from http://www.flickr.com/photos/floridagirl7/4180173503/

Farrell, M. (1966). *Study '66*. from http://www.tcd.ie/artcollections/exhibitions/farrell.php last access 2011-01-04

Flickr. (2011). The Flickr Gallery: Flickr.

Giltrap, C. (2010a). *George Dawson: An Unbiased Eye*. Dublin: Trinity College Dublin.

Giltrap, C. (2010b). Celebrating 50 Years of Modern and Contemporary Art at Trinity. Retrieved Jan 1, 2011, fro m http://www.tcd.ie/artcollections/exhibitions/50years.php

Halliday, J. (2010). RIM announces Playbook tablet device. Retrieved Sep 30, 2010, from http://www.guardian.co.uk/technology/blog/2010/sep/28/rim-playbook-tablet-computer

Ivanov, S. (1875-1951). *Autumn Leaves are Falling*, circa 1906. Retrieved Jan 10, 2011, from http://www.nationalartgallerybg.org/index.php?l=60&id=21

Kaywa. (2011). from http://qrcode.kaywa.com/ last access 2011-01-04

Mac an Airchinnigh, M., & Strong, G. (2010). Social networks and the national art gallery. In T. Hedlund & Y. Tonta (Eds.), *ElPub 2010, Publishing in the networked world: transforming the nature of communication.* (pp. 217-233). Helsinki, Finland: Hanken School of Economics.

Marincola, P. (2006). *What makes a great exhibition?* Philadelphia, PA and Chicago, IL: Philadelphia Exhibitions Initiative. Distributed for Reaktion Books in the USA and Canada by the University of Chicago Press.

Museo Nacional del Prado. (2010). *2010* (March 23).

National Art Gallery Sofia. (2008). Retrieved January 18, 2010, from http://www.nationalartgallerybg.org/

National Gallery of Ireland. (2010). Retrieved January 20, 2010, from http://www.nationalgallery.ie/

National Gallery of Ireland. (2011). Retrieved January 16, 2011, from http://www.nationalgallery.ie/

National Gallery of Ireland. *MyCollections*. (2011). Retrieved June 3, 2011, from http://onlinecollection.nationalgallery.ie/mycollections/view?t:state:flow=a905a902-3e39-4179-94cb-66aeb8865e66.

Newgrange Winter Solstice (2011). [cited 2011 Jan 24]; Retrieved January 20, 2011, from http://www.newgrange.com/winter_solstice.htm

Orela, M. (2011a). ElPub2011 Mobile Digital Art Exhibition. Retrieved Jan 11, 2011, from http://www.flickr.com/photos/mihal_orela/galleries/72157625665088475/

Orela, M. (2011b). ElPub2011 National Gallery of Ireland. Retrieved Jan16, 2011, from http://www.flickr.com/photos/mihal_orela/galleries/72157625713817097/

Orela, M. (2011c). Flickr Gallery Танцьорка.

Orela, M. (2011d). WomanInRed QRtagged. Retrieved Jan 9, 2011, from http://www.flickr.com/photos/mihalorel/5340241080/

Pamuk, O. (2005). *Istanbul: memories of a city*. London: Faber and Faber.

Scott, P. (1974). *Kite!* , from http://www.tcd.ie/artcollections/art-campus/scott-kite.php last access 2011-01-04

Stadtmagazin Echo Münster. (2011). Das Abenteuer der Moderne: George Braque hat es mitgeprägt. Retrieved Jan 12, 2011, from http://www.echo-muenster.de/node/37757

Storr, R. (2006). Show and Tell (pp. 14-31).

Thorpe, V. (2011). 'Gallery rage' mars the Tate's record-breaking Gauguin show. Retrieved Jan 16, 2011, from http://www.guardian.co.uk/artanddesign/2011/jan/16/gauguin-tate-modern-crowds

Trinity College Dublin. (2010). College Art Hire Scheme. Retrieved Jan 14, 2011, from http://www.tcd.ie/Communications/news/news.php?headerID=1570&vs_date=2010-10-1

Where is the Museum of Innocence? (2011). Retrieved Jan 17, 2011, from http://istanbul2010culture.com/2011/01/where-is-the-museum-of-innocence/

Wikipedia. (2011a) H.264/MPEG-4 AVC. Retrieved April 3, 2011, from http://en.wikipedia.org/wiki/H.264/MPEG-4_AVC

Wikipedia. (2011b) Advanced audio coding. Retrieved April 3, 2011, from http://en.wikipedia.org/wiki/Advanced_Audio_Coding

Wikipedia. (2011c) JPEG 2000. Retrieved April 3, 2011

Wikipedia. (2011d). William Shakespeare. The Seven Ages of Man. Retrieved Jan 18, 2011, from http://www.bintmagazine.com/bint_stories/500.php?story_id=475

Wikipedia. (2011e). All the world's a stage. Retrieved Jan 18, 2011, from http://en.wikipedia.org/wiki/All_the_world's_a_stage

Абрашев, Н. (1897 – 1962). Композиция - женска фигура III, около 1950. Retrieved April 3, 2011, from http://www.nationalartgallerybg.org/index.php?l=55&id=3

Иванов, С. (1875-1951). Капят есенни листа, към 1906 г. Retrieved April 3, 2011, from http://www.nationalartgallerybg.org/index.php?l=55&id=21

Always On: Museums in the Mobile Communication Era

Licia Calvi

NHTV Breda University of Applied Sciences, Academy for Digital Entertainment, P.O. Box 391, 4800 DX Breda, The Netherlands. Email: calvi.l@nhtv.nl.

Maria Cassella

University of Turin, Via Po, 17, 10124 Torino, Italy. Email: maria.cassella@unito.it

Abstract: *The paper investigates the use of mobile tools by museums in order to provide mobile access to their permanent collections and special exhibitions. In fact, it deals with the wider topic of how museums tackle the complex issue of communicating with their present and potential audience using modern (i.e., mobile in this case) technologies. The paper presents and discusses the results of a survey that was proposed to Dutch and Flemish museums mainly dealing with modern and contemporary art or with science and technology. We tried to derive some trends and best practices in order to identify a good way to provide an engaging (mobile) experience to museum visitors. These results, although not always stirring in terms of answer percentages and of what most museums seem to be doing with new media, do show a clear interest towards mobile technologies and openness to innovation in the Dutch cultural sector.*

Keywords: *Mobile access to digital cultural heritage; mobile museums; mobile user experience.*

Introduction

For a number of years, we have been investigating the use and the impact of social media on cultural heritage. In particular, we have been studying the use of Facebook by academic libraries (Calvi et al., 2010) and its intrinsic possibilities as a medium to attract new users, to help libraries build and maintain a well-defined and recognised role (Cassella, 2010) and to transfer efficiently their knowledge to their audience (Cassella & Calvi, forthcoming). However, librarians are not the only ones within the cultural sector to experiment with the use of the social tools of the Web 2.0 to advocate, promote, and raise awareness about their collections and services (Boost, 2009; Boost & Calvi, 2009).

This issue is indeed part of the wider concern of cultural institutions to reposition themselves in the vast digital environment and to redefine their role and expertise in a more complex informational context, where users have mostly become remote and information retrieval and discovery tools have improved the strategies adopted to promote cultural communication, to attract new users and to facilitate access to their information.

In this paper, we analyse how museums make use of the mobile user experience by performing an empirical analysis of current practices in museums in the Dutch speaking world (i.e., The Netherlands and Belgium). The museums that we have selected include science and technology museums, graphic design museums and modern and contemporary art museums, i.e., all museums that deal with contemporary art forms and that for this reason we consider more inclined to experiment with modern technology. Apart from the clear design issues that are implicit in its adoption, and which imply not just technological developments but also the (new) ways to enhance users' participation and to promote social inclusion, there is also a more strategic and communicative aspect in choosing to resort to mobile media to promote museum collections. This is why we have been mainly concerned with the question of how museums can use mobile tools to promote access to digital culture, i.e., to promote their exhibitions and permanent collections. We have also highlighted the strong and the weak points in using mobile communication for museums. In addition, we have tackled the issue of how to enhance the mobile user experience for those museums that do already use mobile tools.

Background

Driven by the Apple devices' success,[1] mobile communication is fast spreading all over the world. According to the last Pew & American Life Project Report, 47% of American adults state they access local news on their cell phones or tablet computers. The demographic characteristics of mobile information consumers show that adults using mobile information are very young, affluent, educated and live in non-rural communities. By 2020, mobile devices will be the primary connection tool to the Internet for most people in the world.

The potentiality of mobile technology is enormous. At the same time, mobile communication poses new challenges; for example, it changes people's perception of space and time and the way people look at and experience reality. Above all, as Dempsey (2009) states:

> a discussion about mobile communications, especially when seen in the context of the broader diffusion of network communications, soon touches on many fundamental issues: pedagogy and cognitive approaches; organization and structures; lifestyles and social preferences.

In times of economic crisis, cultural institutions (i.e., libraries, museums and archives) have taken advantage of the huge possibility offered by mobile technology to reposition themselves in the digital environment and to redefine their role and expertise in a more complex informational context, where users have become mainly remote users and information retrieval and discovery tools have improved the strategies adopted to promote cultural communications, to attract new users and to facilitate access to the information they need.

For many years, art museums have been offering their visitors learning experiences that extend beyond traditional exhibit labels with in-gallery kiosks and audio guides. This experience is now moving fast towards platform-specific applications in an effort to capitalize on the commonly-owned portable devices – iPods, MP3 players, Blackberries, smartphones, iPhones, iPads – that visitors already carry in their pockets.

Surveys and Case Studies

From 2005 on,[2] to enhance the users' mobile experience, museums have carried out manifold surveys among visitors to get feedback and comments on mobile technology and services they could offer. In 2008, for instance, the Whitney Museum of American Art performed a survey on visiting users to assess their use of technology to interact with the museum. Among other results, the survey highlighted the visitors' preference to download audio tours from the Whitney's website (74%). Of those who were interested in this, a majority (59%, n=90) would prefer to download the content prior to the exhibit, 25% (n=38) would prefer to download it after visiting the exhibit, and only 16% (n=24) would prefer to download it at the museum. More than half (53%) of the respondents owned an iPod, 14% owned an MP3 player, and only 7% of them owned an iPhone.

Among visitors who participated in the survey, half (50%, n=103) would want to use their personal device for video or information about the museum or the exhibit; 27% (n=56) would not want to, and 23% (n=48) were not sure. Fifty-nine percent (n=115) of them would prefer a combination of video and audio content. Audio and video tours were among the first kind of mobile applications offered to visitors by museums.[3] Initially, mobile applications were mainly supported by traditional players rented at the entrance of the museum.

---

[1] According to the Burson-Marsteller blog (September 2010), the initial three-year growth rate of Apple's iPhone/iPod Touch, launched in June 2007, is ten times faster than the initial three-year growth rate of the online service American Online (AOL).

[2] In 2005, one of the first surveys on the use of technology in museums was the study conducted on a hand-held device developed for the Rembrandt's late religious portraits exhibition held by the J. Paul Getty Museum in Los Angeles.

[3] According to a 2010 Reach Advisors' survey on the use of technology in exhibitions, audio tours, videos and movies are mainly preferred by older visitors. Reach Advisor is an American research firm performing surveys on consumers (http://reachadvisors.com/).

In 2011, mobile technology in museums is moving on and visitors access mobile services from their personal smartphones, tablets, etc., i.e., from any kind of mobile personal device. Sometimes, they need to download a specific application that enables them to access only the information that is relevant to their profile (and interests) as specified, generally online (see below for a short discussion on what the Dordrechts Museum has started to do in 2011 in this respect), prior to the physical visit to the museum.

As a matter of fact, museums are now investigating whether it is still worth renting mobile devices to visitors: in June 2010, a short survey on visitors' mobile preferences was performed by the Indianapolis Museum of Art (IMA). The results indicated that the majority of the people surveyed prefer to access mobile content from their own devices, followed by a fewer number of respondents who would rather rent a mobile device from the museum and by a smaller portion of respondents who prefer to sit and watch a museum video at home.

As for the content visitors prefer to access, 90% answered that they prefer to hear an explanation of the work exhibited from the concerned artist, 83% want to have explanations behind the scenes, 54% prefer to hear comments from the experts, 51% want to access high resolution images, 27% state their interest is in games.[4]

Other surveys have focused on museums to assess their experience when going mobile. A survey on museums conducted in 2009 by the Center for History and New Media (CHNM) at the George Mason University, for example, found that 67% of respondents had implemented or were in the process of implementing a mobile content delivery project. Sixty-one percent of the respondents that implemented a mobile project offered podcasts, 54.5% offered cell phone tours, 36.4% offered an iPhone/iPod touch application.[5]

In January 2011, the second Museums and Mobile annual survey containing data about 738 museums was published. The survey – performed from September 2010 to November 2010 – tried to assess the objectives of museums going mobile, the main challenges in delivering mobile applications and the future perspectives of mobile technology in museums. Thirty percent of the museums surveyed offered mobile interpretation tools to visitors, 23% were planning to do so. This survey highlighted a correlation between the museum annual attendance and the use of mobile technology. Half of the institutions with over 250,000 annual visitors currently use mobile tools compared to less than 20% for those institutions with fewer than 50,000 visitors. Another correlation was among the number of staff working on mobile projects and the offer of mobile tools to visitors. Over two-thirds of the institutions with between one and five members of staff working on digital technologies used or were planning to use mobile tools.

Four most important objectives were identified by museums offering mobile experience:

- To provide supplementary information to visitors;
- To diversify the museum's offering to visitors;
- To engage visitors As part of the museum's experimentation; and
- To create a more interactive experience.

The most important challenges were identified as costs, keeping the content up-to-date and the technical development of the mobile tool.

Methodology

We developed a survey consisting of 27 questions[6] and published it online in March 2011. The link to this survey was sent to about 75 Dutch and 15 Flemish museums, most of which were applied art and modern or contemporary art museums and some science and technology museums. From the survey logs, we identified the profile of the museums that filled out the questionnaire as being graphic

[4] Short results of the IMA survey can be found at the URL: http://www.imamuseum.org/blog/2010/06/22/have-it-your-way-results-from-our-2-minute-mobile-survey/

[5] Museums and Mobile Adoption Survey, Center for History and New Media, open from January, 31st till April, 1st 2009. Survey results are available from Survey Monkey: http://www.surveymonkey.com/sr.aspx?sm=mpcrjZYeT32X_2fVFtptvGGdO3OZb4eP9azwlmPd8uDew_3d

[6] http://pws5.parantion.nl/index.php

design museums, museums of contemporary or modern art and history or architecture museums with an attendance ranging between 20.000 and 270.000 visitors per year.

The survey was also published on an online Dutch museum forum[7] and the post was read by 148 of its members. The survey was accessed by a total of 53 respondents over a period of 10 days. Of them, 14 did not fill it out completely while 9 completed it fully. However, because of the former group of respondents who filled out some parts of it, we have been able to collect more answers for the questions in the first part of the survey. Finally, a couple of respondents sent their reply by email as they could not fill out the survey online completely.

The questions covered rather heterogeneous issues ranging from the reason(s) why the museum had decided to adopt mobile technology (for example, to promote the museum collections and exhibitions; to attract new visitors and increase the number of visitors; to provide better access for visitors; to pilot innovative services; to generate new revenues, etc.), to the kind of services and plug-ins offered via the mobile platform adopted (like a mobile website of the museum; mobile video tours of the museum; mobile video tours of the special exhibitions; e-books and catalogues on permanent and special collections, for instance); from the technicalities involved in the service offered (i.e., does this mobile service rely on visitor's personal device or is the device provided by the museum itself?), to the business models underlying this choice (i.e., do visitors need to pay a fee to be able to use this service or not?). Nevertheless, the core of the questions concerned the museums' general approach to mobile technology, identifying strong and weak points related to this adoption, and the overall positioning of the museum with respect to this adoption.

In the next sections, we will discuss in more detail some of the figures drawn from this survey by clustering them into three categories: services provided, technicalities and underlying business models and museums' general approach to mobile technology.

Mobile Service Provided

Of the 20 respondents who answered the questions on the use of mobile technology by their museums, 55% (11 museums) replied that their museum had not adopted so far any mobile technology whatsoever (and 5% or 1 museum replied that this happens seldom), although a total of 77% indicated that there were plans to develop this service in the future. The services that will most likely be implemented include: iPod Touch/iPhone application (57%), cell phone tour (36%), Podcast (21%), PDA or hand-held self-guides (29%), Android OS (Google) application (14%), mobile style sheet (CSS) making the current website more readable on mobile devices (14%).

However, of the remaining 40% (9 museums) that did use mobile technology, 56% (n=5) claim they did it to promote the museum collections and exhibitions; 44% (n=4) to attract new visitors and increase the number of visitors; 33% (n= 3) to reach potential visitors and provide better access to them; 22% (n=2) to generate new revenues; and 11% (n=1) to pilot innovative services;. Additional replies included: to attract the attention of media and of policy makers (22%, n=2), and to keep up with current trends (22%, n=2). See Table 1 for services and plug-ins provided by museums.

Table 1. Services and plug-ins provided by the museums

4. What kind of services, plug-ins, etc. are you offering via mobile platform? (N_{Total} = 20), (N_{Valid} = 12)	
A mobile website of the museum	33%
Mobile video tours of the museum	25%
Mobile video tours of the special exhibitions	17%
Mobile audio tours (podcasts)	33%
E-books and catalogues on permanent and special collections	8%
Multimedia guides and instructions	42%

[7] http://www.museumserver.nl/

Technicalities and Business Models

Fifty-eight percent (n=7) of the respondents indicate their museum provides mobile devices to the visitors who want to take advantage of this service and 83% (n=10) offer it without asking for a fee. For 75% of the respondents (n=6), the underlying business model therefore relies on free use of the mobile service. No respondents indicated use of a mixed model (i.e., with some free services and some requiring a payment), not the so-called "freemium" (DaPonte, 2010) whereby visitors get free access and the possibility to download the applications they require for free so they are inclined to buy more if they have a good and valuable user experience.

Although the mobile application was not developed by the museum staff (67%, n=8), the museum staff is aware of the resources or tools that were consulted to develop it (as listed in Table 2).

Table 2. Resources used to design the mobile application

13. How was it designed? Please indicate what tools and resources were consulted to design the mobile application (N_{Total} = 20), (N_{Valid} = 12)	
Audience research and evaluation	8%
Content creation	58%
Programming and technical development	25%
User interface design	25%
Don't know: it was developed externally	33%

As indicated in Table 2, museum staff was mainly involved in the creation of the content for the mobile application (58%, n=7), whereas the technical development, here including also users' need analysis, interface design and evaluation have only partially been performed by the museum staff itself (25%, n= 3). Implementing a mobile service has not been a smooth process for 57% of the respondents (n=4), who encountered mainly technical problems (40%, n=2).

The remaining 40% (n=8) who do not offer a mobile service to their visitors indicate as reasons for this choice the ones listed in Table 3.

Table 3. Reasons for not providing a mobile service

11. If you do not provide mobile content, what are the main reason(s) your museum has not? Check all that apply. (N_{Total} = 20), (N_{Valid} = 13)	
Cost	62%
Technical expertise	15%
Staff time	23%
No interest	15%
Lack of institutional support	8%

Museums' General Approach to Mobile Technology

When it comes to identifying the strong points in using mobile museum applications, 33% of the respondents (n=3) indicated that it enables them to communicate better, faster and continuously with users, especially with teenagers; 33% (n=3) that it offers the possibility to reach out to new potential visitors; 33% (n=3) that it offers the possibility to pilot and develop innovative services; 33% (n=3) that it helps users find information and maps of the museum "on the fly" and can increase museum visitors; 22% (n=2) that it is an alternative channel to promote collections.

The most important weak points that were identified by the respondents in the adoption of mobile applications in museums are that mobile technology projects can be costly to implement (11%, n=1); that senior users may be put off and intimidated by mobile technology (22%, n=2); that museum staff has often limited technical knowledge and is not keen on collaborating for mobile projects (11%, n=1); that museum staff has no time to dedicate to mobile technology projects (22%, n=2).

The respondents indicated several ways of tracking or monitoring user behaviour. Some monitor visitors' use of the museum website using Google Analytics; some track user behaviour through the iPhone store and Android market; still others use written surveys that are handed out to visitors after they have used the mobile application (i.e., a PDA in this case) provided by the museum; finally, the majority of the respondents (57%, n=4) indicated that they are not yet monitoring user behaviour nor developing any procedure to do so. Because of the lack of a systematic monitoring protocol, the respondents could not indicate either the users' reactions to the introduction or adoption of a mobile platform by the museum. Only one respondent (i.e., the one using written evaluation questionnaires after usage) replied that users are rather satisfied with the mobile application provided by the museum although some technical problems related to orientation[8] and due to the GPS's not reacting fast enough arose with the use of the PDA. In this case, the user group that was targeted mainly consisted of young users (about 15 years of age). In order to reach out to them appropriately, the content for the mobile application had also been created by young users.

The few museums that are already using mobile applications feel that they had achieved the goals they set up and are already working on an improvement of the application currently under use (more specifically, a second version of the PDA based upon users' evaluation, or the development of new routes inside the museum to be followed with the mobile application).

Discussion of the Results

We started this analysis with the goal of identifying the practices and attitude towards the use of mobile technologies in museums. We decided to focus on the Dutch speaking world, approaching mainly museums in The Netherlands and a few in Flanders, since most information already available concerns the English speaking cultural sector and does not cover cross-border realities.

We were surprised by the limited number of reactions we received, despite the positive attitude most respondents showed and the trust and faith they seem to have in the use of mobile technologies and in the necessity for museums to move in that direction if they want to survive the current changes and evolve together with their audience. One of the most advanced examples of this approach is represented by the Dordrechts Museum.[9] They apply the idea of making the museum visit personal by asking visitors to build their own profile online before starting the actual visit to the museum itself. Once in, via an iPod or their own smartphones, visitors can retrieve information that is adapted to the profile they had created.[10] The Dordrechts Museum combines personalisation and information mobility to enhance the visitor's experience and, in fact, to provide visitors with an experience, *tout court*. This is a reflection of the notion of "experience economy" introduced at the end of the 1990s by Pine and Gilmore (1999) whereby what the visitor (or customer) remembers after a certain event has taken place, i.e., the experience, is what is most valuable to the visitor and what gives importance to the event itself. As Jason DaPonte puts it, "make institutions *addictive*" (2010). This is why many design companies are focusing more and more on experience design[11] and cultural institutions try to exploit this trend as well in order to reach out to new visitors and strengthen the bond with their existing ones. This notion of the experience, borrowed from the experience economy principle and reflected in experience design, can somehow be linked as well to the "four R's" model developed by Prensky (2001a; 2001b)[12], if we assume that every experience, although not explicitly a learning one, does in fact tacitly create and affect learning.

[8] This is an open air museum located in a park.
[9] http://www.dordrechtsmuseum.nl/
[10] http://www.dordrechtsmuseum.nl/over-mijn-dordrechts-museum#
[11] For example, http://www.northernlight.nl/
[12] Namely, Record, Recall, Relate and Reinterpret, where learners collect information; recall data and other resources available; relate them to their social and information network; and process it to build new knowledge.

Conclusion and Future Work

For museums, the mobile experience is a relevant challenge. According to our survey, 70% of respondents who had not so far adopted any mobile technology are going to plan and develop mobile applications soon. However, in projecting mobile platforms and tools for visitors, museums have to tackle manifold issues, namely:

- Design and usability issues;
- Context of use;
- Plethora of mobile CMS platforms;
- Content and tools (iPod/iPhone apps, Android apps, podcasts, video tours, etc.); and
- Business models and sustainability.

Design and Usability Issues

Due to their non-stationary nature, and because of their normal use in dynamic environments, mobile applications do indeed raise specific questions when dealing with the definition of their usability and user-friendliness and require the adoption of specific usability principles when designing their interface (Brewster, 2002), namely:

- The interface has to be designed for a small screen so as to reduce dramatically the amount of information that is visible at a time. More time has to be spent on scrolling and navigating as a result of this;
- The screen resolution has to be lower, which results in a poorer performance in information retrieval tasks;
- The input mechanisms should be limited and may not always be easy to use; and
- Mobile devices do not have fast Internet connections although new generations of technology are currently starting to spread. Networks provide limited coverage and switching seamlessly between different networks is an additional problem. Computational resources like memory size and processor speed are limited as is the autonomy of the devices.

Context of Use

As Raptis, Tselios & Avouris (2005) highlight, museums are a representative example in which the context influences interaction. During a museum visit, visitors interact with the exhibits through mobile devices. They argue that effective interaction design therefore needs to take into consideration multiple dimensions of the context. Mobile projects in museums should take into account these multiple dimensions by continuously integrating the visitor's real world experience and their mobile experience. This can be a rather easy task to tackle when audio tours are used, but it becomes a more difficult task in the case of video tours and games, for example.

Moreover, the context of use of mobile technology is continuously changing. People tend to interact in small and focused chunks of activities, with a reduced attention because of what is happening around them or simply because of being on the move. This may become a critical issue for museums and should be taken into account when designing mobile applications. For example, focus groups findings at the San Francisco MOMA have highlighted that visitors love a short stop in the narration (or video tour) followed by a "Go deeper" option (Samis, 2010).

Plethora of Mobile CMS Platforms

Designing user interfaces and tools for mobile devices in museums is a huge challenge considering the wide heterogeneity of operating systems, technologies, devices and applications potentially available. Samis (2010) argues that device and distribution infrastructures should be conceived in order to ensure equal access to information for all visitors, tech-savvy persons or not. Museums should also adopt a single mobile CMS that publishes to multiple customized interfaces: in-house multimedia, iPhone, iPad apps, Android apps, or any other platforms.

Content and Tools

The content to be mobilized should be very carefully selected in order both to enhance visitors' experience and to keep costs under control. A selection of the tools to adopt to mobilize museum collections and exhibitions is another relevant issue to consider. Visitors' demographic characteristics should be carefully investigated. Games, for example, are traditionally more effective in educational

projects for a strong interaction with young children and families, but they are becoming increasingly valuable also for young people as the distinction between games and utilities, private spaces and public spaces is blurring, while video/audio tours and movies can be particularly appreciated by senior visitors.

Business Models and Sustainability

According to our survey, costs are a wide concern and may hinder or prevent a wider adoption of mobile technology in museums. Therefore, business models should be carefully planned in order to guarantee sustainability to museum mobile projects. Start-up costs are a big issue but also content update and software maintenance can be costly. Sustaining the whole system is a challenge. Apps do not usually generate income for developers and an iPhone app may cost up to 35,000 dollars to develop (Proctor, 2010).

Although in our survey we found no evidence of mixed business models, we suggest that in order both to support mobile technology costs and to accomplish the museums' mission to offer broad access to collections, mixed business models should be adopted. For example, mobile tools for enhancing access to permanent collections could be offered to visitors free of charge while mobile tools for special exhibitions could be charged or vice versa. The freemium approach mentioned earlier (DaPonte, 2010) is also a possible solution to this.

All these concerns have an impact on the decision to provide mobile access to museum collections and to support a mobile experience to museum visitors. However, if they want to be always on, museums have to experiment with mobile applications and social tools. As future work, we would like to make a comparative study and investigate the adoption of mobile applications in museums in two European countries, like The Netherlands and the United Kingdom.

Acknowledgements

We wish to thank all the respondents to our survey.

References

Boost, A. (2008/2009). *Digitale culturele communicatie: de rol van sociale media in de communicatie tussen culturele instellingen en hun doelpubliek (Digital cultural communication: the role of social media in the communication between cultural institutions and their audience).* M.A. thesis in Journalism, Lessius College.

Boost, A., & Calvi, L. (2009). *Digital cultural communication: the role of social media to enhance the communication between cultural institutions and their audiences*. Paper presented at the Digital Strategies for Heritage Conference (DISH '09), Rotterdam, The Netherlands.

Brewster, S.A. (2002). Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing* 6, 188-205.

Calvi, L., Cassella, M, & Nuijten, K.C.M. (2010). Enhancing users' experience: a content analysis of 12 university libraries Facebook profiles. In T. Hedlund & Y. Tonta (Eds.), *Publishing in the Networked World: Transforming the Nature of Communication. 14th International Conference on Electronic Publishing 16-18 June 2010, Helsinki, Finland* (pp. 258-269). Helsinki: Hanken School of Economics. Retrieved April 7, 2011 from https://helda.helsinki.fi/handle/10227/599.

Cassella. M. (2010). Comunicare con gli utenti: Facebook nella biblioteca accademica. *Biblioteche oggi, 28,* 3-12.

Cassella, M., & Calvi, L. (Forthcoming). *The effect of social software for academic libraries*. IGI Press.

Center for History and New Media. (2009). *Museums and Mobile Adoption Survey*. Retrieved April 7, 2011 from http://www.surveymonkey.com/sr.aspx?sm=mpcrjZYeT32X_2fVFtptvGGdO3OZb4eP9azwlmPd8uDew_3d

DaPonte, J. (2010). Jason DaPonte keynote. Retrieved April 7, 2011 from http://tatehandheldconference.pbworks.com/w/page/28353070/Jason-DaPonte-Keynote

Dempsey L. (2009). Always on: libraries in a world of permanent connectivity. *First Monday*, *14*. Retrieved April 7, 2011 from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2291/2070.

Pine, J. & Gilmore, J. (1999). *The Experience Economy*. Boston, USA: Harvard Business School Press.

Pocket-Proof & Learning Times. (2011). *Museums and mobile annual survey*. Retrieved April 7, 2011 from http://www.museums-mobile.org/survey/

Purcell K., Lee, R., Rosenstiel T., & Mitchell, A. (2011). *How mobile devices are changing community information environment.* Retrieved April 7, 2011 from http://www.pewinternet.org/~/media//Files/Reports/2011/PIP-Local%20mobile%20survey.pdf

Prensky, M. (2001a). *Digital natives, digital immigrants, part I*. London, UK: MCB University Press.

Prensky, M. (2001b). *Digital natives, digital immigrants*, *part II : Do they really think differently?* London, UK: MCB University Press.

Proctor, N. (2010). *Mobile social media and emerging business models*, keynote presentation, Tate Handheld Conference (THC '10). Retrieved April 7, 2011 from http://tatehandheldconference.pbworks.com/w/page/29018648/Keynote-presentation:-Nancy-Proctor-2010.

Raptis, D., Tselios, N., & Avouris, N. (2005). Context-based design of mobile applications for museums: a survey of existing practices. In: *Proceedings of the 7th International Conference on Human Computer Interactions with mobile devices & services (MobileHCI '05).*

Samis, P. (2010), *Culture shifts: museums and mobiles in the age of social media*, keynote presentation, Tate Handheld Conference (THC '10). Retrieved April 7, 2011 from http://tatehandheldconference.pbworks.com/w/page/29523573/Keynote-presentation:-Peter-Samis-2010.

Whitney Museum of American Art. (2008). *Audio Guide technologies survey final report*. Retrieved April 7, 2011 from http://tatehandheldconference.pbworks.com/f/Whitney+Final+Report-+Appendices+G-I-revised.pdf.

How Should We Read New Media and New Technologies?

Erencan Gökçek

MA Student in Digital Media: Technology & Cultural Form at Goldsmiths, University of London, UK.
E-mail: erencang@gmail.com

Abstract**:** *According to conventional wisdom we frequently hear that "new media" and new technologies are revolutionising the world we live in. By investigating these presumptions in this article, I am going to try to review different critical perspectives interpreting new media and question the distinction between new and old media through Bergson's philosophical method in order to identify "false problems" and "false divisions" as well as different concepts that we encounter in the world of new media today. I will also investigate what's new about new media by considering their contiguity with old media as well as re-evaluating the political connotations of terms such as "convergence" and "interactivity" which are regarded by many commentators as false problems within new media. In this article I shall explain why we should use the following terms: "remediation" and "intra-activity" as opposed to "convergence" and "interactivity", respectively.*

Keywords: *New media; digital media; convergence; interactivity; intra-activity; remediation; intermediation; repurposing; digitization.*

Introduction

The question of technology and its use has been the main focus of media studies from the 1960s onwards, especially when it comes to "new media studies" and new technologies. It was McLuhan's idea that "the content of any medium is another medium" (McLuhan, 1968, pp. 15-16). Logical though this may seem, McLuhan goes on to embrace a rather technologically deterministic approach. Focusing on this conception of media in their book *Remediation: Understanding New Media* (1999), Jay Bolter and Richard Grusin explain how they came up with their concept of "Remediation": "We can reject McLuhan's determinism and still appreciate his analysis of the remediating power of various media" (Bolter & Grusin, 1999, p. 77).

However, before considering their purpose of focusing on media *as dynamic phenomena,* we have to understand the main problems of analysing or reviewing the ways in which Marshall McLuhan and Raymond Williams (1975) referred to media and how they (or indeed whether they) received this concept of *dynamic phenomena.* We should also refer to other arguments as promulgated by Manovich in his book *The Language of New Media* (Manovich, 2001) and how he perceives new media and distinguishes them from other types of media. We should try to understand the media not just as one "object" but as a "collection of objects" which should be analysed economically, socially, culturally, politically, philosophically, theoretically and technologically.

Therefore, they need to be seen not as a separate phenomenon existing outside the world we currently live in but rather as a complex, hybrid process, which Bolter and Grusin (1999) call "remediation" that is also similar to Lister *et al*.'s approach in their book *New Media: A Critical Introduction* (Lister, et al., 2009) as well as Hayles' term "intermediation" in her book called *My Mother Was a Computer* (Hayles, 2005). We should also acknowledge Karen Barad (2007) who comes up with the idea of "intra-action" which will be elucidated later in this essay.

I would argue that it is not helpful for us to understand the media by singling out one individual aspect at the expense of others, like Marshall McLuhan and Raymond Williams (the former being too *technicist* and the latter too *humanist*). Therefore, I suggest that, in order to understand the media, whether it be analogue, digital or new media, it is best not to be reductionist, as it has always been too easy to fall into the trap of saying it is either a "revolution" or "business as usual".

There have been tendencies to depict new media as analogous to computing or a giant computer in accordance with Manovich: "From media studies, we move to something that can be called 'software studies' – from media theory to software theory" (Manovich, 2001, p. 48) However, it is also argued

that this perception of "new media are becoming a computer/computing" is simply misleading and is regarded by many commentators such as Lister *et al* (along with Bolter and Grusin) as being too deterministic and too progressive as it cuts off the continuity that exists between new media and old media and sees them as the last stage in the course of technological as well as cultural progress. It basically suggests that it acts as a substitution, a replacement and a breakthrough. At best, it might be useful in order to perceive what is currently happening; at worst, it is simply wrong as it has unhelpful connotations which suggest a certain view of progress.

My analysis, in terms of understanding the media, will be similar to that of Bolter and Grusin who put forward the term "remediation" on the basis of understanding the media as *dynamic phenomena* (Bolter & Grusin, 1999, p. 44). I will also suggest the term "intermediation" used by Hayles (Hayles, 2005, p. 15), as well as referring to Lister *et al*. who help us to understand "what is new about new media" better in non-linear terms (Lister *et al*., 2009, p. 9).

What are 'New Media'?

Despite the ideological connotations and conceptual limitations of the "newness" as it may be evaluated as better compared to the old, this term can arguably be inclusive and useful as it is not reducible to any kind of existing technologies as long as we are aware of technological over-determinism. Some terms such as "interactive media" might be one example of unhelpful, reductionist definitions as they imply interactivity which is a rather problematic and questionable concept.

I will revisit the possible reasons why I use the term "new media" later in this essay even though they are not completely new. When it comes to defining new media, interestingly enough, although Manovich and Lister *et al*., along with Bolter and Grusin, seem to be split over this issue, they seem to mention the difficulties of describing exactly what they constitute. Nevertheless, it is Manovich who tries to give us a shortlist of new media in his book *The Language of New Media.* For him, they (generally) consist of "The Internet, websites, computer multimedia, computer games, CD-ROMS and DVD, virtual reality" (Manovich, 2001, p. 19). On the other hand, it doesn't seem be so clear-cut for Lister *et al.,* who consider the main characteristics of new media to be "digital, interactive, hypertextual, virtual, networked and simulated" (Lister *et al.*, 2009, p. 13). For them, new media means:

> Those methods and social practices of communication, representation, and expression that have developed using the digital media, multimedia, *networked* computer and the ways that this machine is held to have transformed work in other media: from books to movies, from telephones to television (Lister *et al.*, 2009, p. 1).

Despite the usefulness of these terms as a means of identifying what new media look like, they don't necessarily tell us what they are. Questioning the newness of new media does not necessarily mean that either "there is nothing new under the sun" (as Lister *et al*. disagree with that as well) or that there has been significant subsequent change or replacement within the media, as Manovich tends to say. The fact is that we haven't seen a change in the way in which computers have swallowed the whole of analogue media; that is simply not the case. However, there is, certainly in part, something new in new media ever since the dramatic recent expansion in the availability of PCs and especially ever-increasing mobile devices, which are the main theme of this year's conference, surrounding users or audiences enabling them to become more engaged with mediation and the medium itself.

Remediation

Remediation, according to Bolter and Grusin, can be defined as: "New media are doing exactly what their predecessors have done: presenting themselves as refashioned and improved versions of other media" (Bolter & Grusin, 1999, p. 15). From this statement one can see the connection and co-existence between new and old media and recognise that their relationship is *cyclical* rather than linear. It's a perpetual cycle that doesn't necessarily progress to a finite conclusion. The term "refashion" absolutely reveals what is involved. Remediation expresses a process through which not only new media but also old media re-fashion, recycle and, last but not least, *repurpose* themselves. This continuity may also mean an ongoing conflict or an "entanglement" between opposing forces

which are brought into being in order to save and balance each other's existence. No single medium today seems to submit to another's authority. Instead, along with the entities surrounding them, they tend not to interact but in fact "intra-act" with each other and are ultimately mutually counterbalanced in order to keep pace with the rapid development of advanced technology.

In order to see how connected and networked the media are today, we should visit Bolter and Grusin's diagnosis of the status quo:

> Digital visual media can best be understood through the ways in which they honour, rival, and revise linear-perspective painting, photography, film, television, and print. No medium today, and certainly no single media event, seems to do its cultural work in isolation from other media, any more than it works in isolation from other social and economic forces (Bolter & Grusin, 1999, p. 15).

If we apply this intriguing interpretation to our daily lives, we can undoubtedly recognise the inevitable co-existence of TV channels, newspapers (whether paper or digital) with even Smartphone, tablet PC applications or computer software resembling each other's content, design, purpose, etc. We can't possibly talk about a complete original platform which works in isolation: be this software, an application, perhaps even mobile phones nowadays. Thus, smartphones not only don't offer themselves *per se*, but also encourage their customers to be a part of network society.

Drawing upon the "double logic of remediation" in their book, Bolter and Grusin reveal the nature of "immediacy" meaning, in a nutshell, realism and "hypermediacy" indicating, as a shorthand definition, fascination with media in themselves. These are the essential elements of the dynamic process that can be observed throughout the history of media. For them, remediation works within the logic of these two terms. *Immediacy,* referring to its connections within the context of media and mediation, expresses the medium's unending desire to disappear in order to create the sense of transparent presence, in other words, abandoning its audience's idea of what is it like to be" in a particular place or position. This could be a computer game trying to achieve a "photorealism", or it could be a cinema film trying to create "authenticity". In short, it raises the tantalising prospect of ultimate interface by eliminating immediacy. In their book, the most extreme example they use in the current context of cinema is "Hollywood to replace actors with computer generated animations". In the final analysis, it posits a hypothesis of modal immediacy that wants to remove itself "so that the user is no longer aware of confronting a medium" (Bolter & Grusin, 1999, p. 25).

The other necessary half of remediation is *hypermediacy*. In their view, it is quite the opposite of immediacy. Hence, they define it as the "double logic of remediation". For example, it could imply displaying lots of TVs on a single computer screen. Quoting William J. Mitchell (1994), they say that it reflects a "process or performance rather than finished artwork." Today, for them, what we have been viewing on our screens from the 1980s onwards represents a transition from "homogeneousness to heterogeneousness": diverse, fragmented and multi-functional screens that "compete for users' attention". This also refers to a shift in the form of representations from text to icons in which command-based computing is replaced by interface-based technology which can also be flexible depending on what users want in terms of their size, number, positions on the screen, etc.

If we are to see how they are connected as well as the double logic they constitute, they propose that, while immediacy wants to make itself invisible, simultaneously, hypermediacy, in other words, a "collage aesthetic of different media forms and styles" makes us aware that it exists as a medium that expects to capture our fascination. They claim that those two elements have been in interplay throughout history (Bolter & Grusin, 1999, p. 31). By combining these two characteristics of remediation, this concept enables us to recognise new media as much as all media to be a cyclical representation of history rather than a linear continuum.

So, if we go back to their quote claiming "No medium today, and certainly no single media event, seems to do its cultural work in isolation from other media" we can see the justification of their claim through events we have witnessed via the media such as Michael Jackson's death, the "credit crunch", "Wikileaks", etc. Clearly, we cannot possibly separate these events from their mediation. Bolter and Grusin argue that there is a complex correlation between an event and its mediation; therefore there isn't necessarily a dominant element. Rather, as mentioned at the beginning of this article, it's a "dynamic, progressive and complex process", which should be approached from a variety of viewpoints. Therefore, it's their *hybridity* which should not be overlooked.

Bergson's Philosophy of False Problems and Divisions

In order to analyse the media in this respect, we should seek to distinguish "false problems" and "divisions" which are stated as the first rule of Bergson's philosophical approach known as "division of reintegration" (Deleuze, 2002). We should briefly remind ourselves of the three rules of this methodology:

1. Distinguishing between true and false problems;
2. Distinguishing between differences in degree and differences in kind;
3. Considering the object of our inquiry in terms of its temporality.

The reason why we refer to this rather theoretical analysis is not that we want to study Bergson's philosophy; instead, we will seek to apply this approach in order to distinguish the false problems that we encounter today. It is useful to pinpoint the real problems in the media that we currently critique in this essay as dynamic *phenomena* and it is helpful for getting rid of the ideas that might not be relevant for understanding what is going on in the world of new media and technology at the moment. It is Bergson who simply wants us to come up with better questions and better phrases. For him, these questions might be more interesting and useful than the answers.

Concurrently, we will review key concepts and ideas, in particular the notion of "convergence" and "interactivity", usually associated with new media which will be regarded as "false problems". Although they might not be entirely inappropriate concepts in the context of different issues and they can tell us about new media to a certain extent, these are terms which might constitute certain political connotations such as "progress" or represent an achievement from old to new, from A to B. They do not help us to understand what contemporary media and technologies actually are.

So, if we offer initial definitions, convergence is generally the "coming together of previously distinct technologies, separate economics and separate cultures." On the other hand, interactivity is "a more powerful sense of user engagement with media texts, a more independent relation to source of knowledge, individualised media use, and greater user choice" as opposed to that of old media, which are usually accused of being "passive" (Lister *et al.,* 2009).

Of course, just as we are distinguishing between false problems or divisions, it will be suggested that *there are better terms in order to understand the media as a hybrid, organic phenomenon.* We should not fall into the trap of "limited dualism" or "binary oppositions" such as "theory and practice", "technology and use", "determinism and constructivism" as they are difficult to avoid and render the debates surrounding the new media redundant. These fields are not *divisible* as it is neither possible nor useful to draw a line or establish a boundary between them. Therefore we should refer to *remediation* instead of *convergence* and *intra-activity* rather than *interactivity*.

What Actually are Convergence and Interactivity?

The reason why convergence is a false problem is because it is associated with technological progress emphasising "space" rather than "time" in terms of Bergson's analysis. As the term *convergence* might suggest a change in the way in which the amalgamation of different technologies is leading to one being replaced by another, *remediation* seems to be a better word implying "process" rather than "progress".

For Katherine Hayles, "convergence" demonstrates the danger of "linear causality" for convergence enthusiasts such as Manovich, who, as we have already explained, has misread the media in that sense. 'The convergence of all media into one" for Hayles, "has the effect of flattening" this process into *teleology*. Using the example of print, she says that digital text has already affected print text in terms of the printed books we are seeing today, "moving from straight text to visual" (Hayles, 2005: 31). In fact, for Lister *et al.,* convergence seems to bring about fragmentation as opposed to the notion of synthesis: "Television, online video delivery, internet communications and telecommunications combined in one 'black box' that will look something like a phone" (2009, p. 202).

However, in the world we currently live in, the notion of convergence implying progress does not stand up to scrutiny. The state of play in the realm of computers and platforms is quite the opposite. It is the apparently endless proliferation of these devices and the rapid dispersal of innovative hybrid technologies which seem to challenge this idea of convergence. A perfect example of this trend is the

Apple tablet computer iPad[1] (which is already becoming obsolete) and its hybrid status in the world of mobile devices. This is far from unique; however, it also begs the question that if the idea of convergence isn't a false problem, then why do we still have laptops, mobile phones & desktop computers which don't seem to be coalescing but are actually becoming fragmented and dispersed?

The emergence of these technologies does not justify the claim of *singularity,* the term that is used in artificial intelligence (AI). On the contrary, they represent the outcome of a process that can be defined as remediation, reconfiguring existing tools and concepts whereby new media re-fashion and re-cycle old media technologically and indeed commercially. This can be alternatively characterised as the metastasis of the divers (hybrid) gadgets permeating us day by day as opposed to their being submerged or packaged into a black box. Consequently, it's the users who get caught up within an "upgrade culture" which doesn't end.

However, Henry Jenkins in his 2006 book *Convergence Culture* perceives that there is indeed something that "has converged". It's the content. In his words, "convergence represents a cultural shift. . . it is we who are convergent. . . to seek out information and make connections between them" (Jenkins, 2006, p. 3) Since a merger of all the portals and mechanisms of communications technologies has yet to be seen, the claim of media convergence is not convincing.

There are some other approaches towards the notion of interactivity that might not be a false problem for some commentators such as A. Barry. According to him (2007) interactivity might be an alternative to "discipline". In Foucauldian theory (1991), this comes from "power" in order to establish "correct training" through which discipline can create ideal citizens, in this case "active, neo-liberal citizens". Consequently, in contrast to the old discipline, this new interactivity suggests "you may" and "discover" as opposed to "you must" and "learn". This seems to be similar to the thinking of the Frankfurt school, which criticised the cultural and political implications of broadcast media and now seems to be celebrating the emergence of new media. This is because for them it is a more imaginative and democratic use of the media as they have long desired (Lister *et al.,* 2009: 74).

In contrast to these arguments, it is clear that the idea of interactivity carries with it the notion of progress, therefore we will suggest the term 'intra-activity' as mentioned by Barad (2007, p. 33). Because interactivity tends to see users and the media as separate entities and regards the media as objects by emphasising space and *spatiality* rather than *time* and *temporality*, it foregrounds the concept of "progress" instead of acknowledged or unacknowledged "process". For her, the relationship between new technologies and users is based on "co-existence" and "co-institution". It is even more dynamic than *interactivity* as it emphasises entanglement whereby users and media "co-evolve". If anything, they go hand in hand, acting and living together. They are part of the same reality and eventually part of us rather than being separate entities (Barad, 2007).

On this matter, Lister *et al*. again seem to have a lot to say. Quoting Aarseth (1997, p. 48) in their book, "To declare a system interactive is to endorse it with magic power", they think that interactivity has ideological meanings. When considering the value added characteristics of new media, interactivity clearly has something to do with current neo-liberal, individualist discourses as has been discussed before. They argue that it prefers to favour "consumerism" and will eventually sustain the concept of "upgrade culture", which treats users as "consumers" who are now enjoying the "freedom of choice", choices that are coming from the "array of possibilities offered by the market". Metaphorically, it can be seen to resemble an ocean of opportunity within which users are floating around and exploring a plethora of possible options.

At this stage, I would argue that these choices mostly point to a commercial end-product. In fact, it indicates a change from "viewer" or "reader" to "user" representing the new media audience. Therefore, interactivity has more to do with consumerism than *democracy*. Artist Sarah Roberts' view might also be enlightening:

---

[1] An iPad2 advertisement narrative goes like this: *"This is what we believe: Technology alone is not enough. Faster, thinner, lighter…Those are all good things. But when technology gets out of the way, everything becomes more delightful, even magical. That's when you leap forward. That's when you end up something like this."* Doesn't it sound like immediacy to you? [The transcript of this advertisement has been taken from the video on Apple's web site. Copyright © 2011 Apple Inc. All rights reserved]

> The illusion that goes along with [interactivity] is of a kind of democracy…that the artist is sharing the power of choice with the viewer, when actually the artist has planned every option that can happen…it's a great deal more complex than if you [the user] hadn't had a sort of choice, but it's all planned (Penny, 1995, p. 64).

Another problem with interactivity is that it has a rather rhetorical approach to old media claiming that these weren't really "active". But if we take into account what Roland Barthes says in *Death of the Author*, he claims that we are never passive readers when we read a book; as far as he is concerned, we, i.e. the readers, are actually writing it at the same time we engage with it (Barthes, 1977). Similarly, in her book *Meeting the Universe Halfway,* Karen Barad says in the prologue that "It is not so much that I have written this book, as it has written me" emphasising the process of *intra-activity* which foregrounds *act with* as opposed to interactivity that implies a less dynamic meaning like act *between*.

In order to avoid being too absorbed by the deeper philosophical issues concerning interactivity and convergence, we will move on from these two terms, so that we can see the media more in terms of "their dynamics" & "power relations" in the context of *time and temporality* rather than *space and spatiality*.

Having established these principles in congruence with this debate on the nature of new media and recognising them to be misleading "false divisions", we have already suggested the terms *remediation* and *intra-activity* in place of *convergence* and *interactivity,* respectively. If we move on to the second rule by taking into account Lister *et al.*'s perception of new media as "continuity" rather than "novelty", it's possible to say that new media are not necessarily entirely new, therefore constituting "difference in degree" rather than "difference in kind". Thirdly and finally, if we proceed to the "re-invention" of the problem by "Stating the problem in terms of time rather than space", we can read the media and mediation from the perspective of *indivisibility*. There is no clear boundary that can be distinguished; it can be viewed as an abstract ongoing hybrid process, but gaining "access" or making "contact", as Bergson would put it, is not easily achievable. It's not a single object that can be reduced to one single aspect, as claimed by the Frankfurt school which regards this process as "interaction". That is to say that the media are not a set of entities existing somewhere outside the cosmos: i.e., a third (or middle) party "object" that sits on top of the social hierarchy as a static phenomenon which might also be seen to be "performative" and responsible for certain events. This is an utterly false interpretation of what we are witnessing today. We should recognise this to be the (re)mediation of a perpetual cycle: the "mediation of mediation". All in all, according to Bergson's logic, we should not dismiss *time* at the expense of *space*.

Why "New Media"?

It is clear that we cannot articulate a definite distinction between old and new media. We cannot separate them. But paradoxically, we still identify them as "new media" without acknowledging their newness which foregrounds an overarching concept and promotes the false notion of "subsuming" the old to imply a sense of "progress". On the contrary: commentators seem to be employing historical tools, which had seemed to have vanished in the context of "old media" and appear to be looking for any opportunity to revive them in the context of digital media. According to Lister *et al.,* there are three main reasons why we continue to be fascinated with new media despite the fact they are not genuinely new:

1) New media are thought of as epochal; whether as cause or effect, they are part of a larger, even global, historic revolution;
2) There is a powerful utopian and positive ideological connotation attached to the concept of something being "new";
3) It is a useful and inclusive "portmanteau" term which avoids reducing "new media" to technical or more specialist (and controversial) terms (Lister et al., 2009, p. 10).

If we briefly review these reasons in sequence, the first suggests a transitional shift, rather than a quantum leap from one era to another: e.g., the evolution from modernism to post- modernism, from apes to *Homo sapiens,* etc. It's a discourse that presumes "everything will be different from now on". So these are again questionable changes that are not entirely useful or totally teleological. Secondly, new media have ideological connotations alongside false promises. They seem to "promise more than they can deliver". Ultimately "they sound utopian" adumbrating false hopes and promises attached

which are commercially attractive because for consumers, "new is always better than old". Finally and most interestingly, new media are defined as such because they are arguably a "less false" option available in contrast with digital media or interactive media which might sound "narrow" and "reductionist".

Using the example of the 1980s ghetto blaster and also the 1940s radiogram in their book, Lister *et al.* (2009) ask us to compare them to the new arrivals, such as iPods, in terms of their use. When we think about those people in the 1980s who were carrying their ghetto blasters on their shoulders, would it be possible to expect the same behaviour from commuters these days? However, what about those commuters who have their iPods plugged into their ears? In this case, aren't new media and new technologies "enabling" or "constraining" analogue media? But does that constitute a complete change? "No" say Lister *et al.* In fact, they are not even different technologically; we cannot even determine a mutually exclusive difference between "new" and "old" media.

The proliferation of mobile devices obviously leads to ubiquitous computing but towards a way in which it creates an "extreme form of hypermediacy" rather than singularity. However, the notions of ubiquitous computing and convergence are not new and remain "utopian and futuristic" as in films such as Apple's *Knowledge Navigator* and Sun's *Starfire* (Bolter & Grusin, 2000, pp. 218-219). In fact, just as the radio has survived television, the web has survived applications or push technologies as opposed to the notion posited in Chris Anderson and Michael Wolff's article "The Web is Dead, Long Live the Internet" which arguably articulates little more than what they suggest in the title (Anderson & Wolff, 2010):

> It seemed just a matter of time before the Web replaced PC application software and reduced operating systems to a "poorly debugged set of device drivers," as Netscape cofounder Marc Andreessen famously said. First Java, then Flash, then Ajax, then HTML5 — increasingly interactive online code — promised to put all apps in the cloud and replace the desktop with the webtop. Open, free, and out of control.

It is clear that nothing is vanishing in favour of convergence as much as they may have imagined; on the contrary, it's not a revolution but a remediation of technologies as well as their content on a larger scale. These more technologically advanced devices as well as their software increase the intensification of mediation and always offer faster, better, more advanced and eventually "more immediate" versions of themselves to be in touch with reality. These may be globally distributed via friends, colleagues, news, journals, books, etc. ceaselessly promoting more and more consumption. Consequently, this "upgrade culture" always expresses the rhetoric of immediacy, i.e., "what is being published right now?" "What's happening now?", and inevitably leads to obsolescence. Unsurprisingly, your mobile device rapidly becomes obsolete; your tablet computer becomes outdated.

Intermediation

Since the focus of this essay is *dynamic phenomena,* it would seem that Katherine Hayles' interpretation of the term *intermediation* deserves to be reviewed. While thanking Bolter and Grusin for suggesting the term *remediation,* in her book *My Mother Was a Computer* she doesn't seem to be satisfied with this signifier as it is insufficient to describe the presence of "feedback loops" and "multiple causality. . .among media". Hayles prefers to use these terms in order to explore complex modal relations: "particularly analogue and digital". She proposes that intermediation is a more suitable word, because it emphasises cycles rather than a "starting point" in contrast to remediation (Hayles, 2005, p. 33). As for its definition, it is basically, in her words, an "endless novelty of combinations" in which digital media recycles print media via constant "connectivity" that is always dynamic and multi-layered. In other words, Hayles is not necessarily only interested in new or old media but also human beings, machines, languages and codes that she thinks also affect this process of intermediation. At this stage, if anything, Hayles remediates or intermediates along the same lines as aforementioned commentators, in her view of the media as dynamic phenomena:

> As the technology becomes more sophisticated, this dream is not fulfilled; instead, embodied subjects are interpolated ever more tightly into a circuit controlled by capitalistic dynamics. As bodies enter more completely into the circuit, subjects cease to be *circumscribed* by the dynamics and are *constructed* through them (Hayles, 2005, p. 9).

Keeping in mind Hayles' entanglement, this new way of looking at the technologies, we can say that technology doesn't necessarily provide a means to an end anymore, as we tended to assume in the past, as for instance Heidegger (1954) thought in his essay *The question concerning technology*. Rather, there is now a new relationship emerging through this co-existence. We live with technology, in fact; we cannot live without technology. We are so engaged and entangled with mobile technologies in particular, we are now talking about post-humans and cyborgs and cyberspaces. Within this milieu we have become part of technology.

The Question Concerning University and Academic Publishing

Drawing on university and open access research, Gary Hall writes in his book *Digitize This Book!* (which has indeed been digitized) that Google's Book Search project aims to scan digitally large numbers of books produced by some Western academic institutions and then make their contents available online for full text searches. The aim is to eventually produce a "universal virtual library" of available knowledge to rival the library constructed at Alexandria around 300 B.C., which was then estimated to contain between 30% and 70% of all known books (Hall, 2008, p. 7). He goes on to say:

> But even if Google were intending to make all of the texts they are digitizing available for free in their entirety, a book still has to be published first in ink-on-paper form for Google to digitize it. This in turn means that, more often than not these days, a book still has to be capable of generating a financial profit for its initial publisher before it can appear as part of the Google Book Search Project (Hall, 2008, pp. 6-7).

All these discussed issues invite us to re-consider our perceptions to date:

First, convergence does not seem to be evident in this era of constant technological development. What that means is that we are less likely to see these mobile technologies converging into one "black box" that looks like a mobile phone. Conversely, they will be more dispersed and fragmented. The consequences of this process can be predicted to be *decentralisation* and *deterritorialisation* of commercial commodities – including of course academic works/research – through the digitization and eventual intensification of new technology consumption. They will always try to perpetuate the rhetoric of innovation, to sell more mobile devices claiming faster, more powerful and more immediate results. However, achieving immediacy in this era still seems to be an impossible aspiration.

Second, as a consequence of this fragmentation, however, technology permeates us: we act with technology. Hence it's intra-activity. It's not a tool standing outside society; on the contrary, it constitutes us. I believe it's critical for academics, students, librarians, executives, entrepreneurs, etc. to re-think our different mechanisms for retrieving information and knowledge through these technologies. It seems that there is no escape from what Castells calls our "network society" as we become more connected to each other (Castells, 1996).

Finally, by endeavouring to analyse what lies behind these technological transformations, I should ask: Are they really changing our lives fundamentally? Are these changes actually significant or superficial? Are they really capable of delivering what they are promising? Remediating Bergson here again, we should ask better questions and eliminate the false problems to distinguish these discrepancies to define not only what new media and technologies are, but also what we are.

Acknowledgements

I would like to thank Sarah Kember for shaping the ideas in the first place, Sebastian Melo for his useful comments and Tamlin Walenn for his endless help and finally special thanks to Yaşar Tonta for his encouragement of this paper.

References

Aarseth, E. (1997). *Cybertext - Experiments in ergodic literature*. Baltimore, MD.: Johns Hopkins University Press.

Anderson. C. & Wolff, M. (2010). The Web is dead. Long live the Internet. *Wired*, September 2010. Retrieved May 29, 2011 from http://www.wired.com/magazine/2010/08/ff_webrip/all/1

Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Durham, NC: Duke University Press.

Barry, A. (2001). *Political machines: Governing a technological society*. London: Athlone Press.

Barthes, R. (1977). The death of the author. In *Image, Music, Text.* Translated by Stephen Heath. Retrieved 13 January 2011 from http://evans-experientialism.freewebspace.com/barthes06.htm.

Bolter, J.D. & Grusin, R. (1999). *Remediation: understanding new media*. Cambridge, MA: The MIT Press.

Castells, M. (1996). *The Rise of the Network Society*, Oxford: Blackwell.

Deleuze, G. (2002). Intuition as method. In *Bergsonism*, trans. Hugh Tomlinson and Barbara Habberjam, New York: Zone Books.

Foucault, M. (1991). *Discipline and punish: The birth of the prison*. Translated by A. Sheridan. London: Penguin.

Hall, G. (2008). *Digitize this book!: The politics of new media, or why we need Open Access now*. Minnesota: University of Minnesota Press.

Heidegger, M. (1954). *The question concerning technology*. Retrieved April 4, 2011 from http://www.lcc.gatech.edu/~xinwei/classes/readings/Heidegger/heidegger_techquestion2.pdf.

Hayles, K. (2005). *My mother was a computer: digital subjects and literary texts,* Chicago: The University of Chicago Press.

Jenkins, H. (2006b). *Convergence culture: Where old and new media collide.* New York: New York University Press.

Lister, M., Dovey, J., Giddings, S., Grant, I., & Kelly, K. (Eds.), (2009). *New media: A critical introduction*. 2nd ed. London: Routledge.

Manovich, L. (2001). *The language of new media*, Cambridge, MA: The MIT Press.

McLuhan, M. (1968). *Understanding Media*. London: Sphere.

Mitchell, W.J. (1994). *The reconfigured eye: Visual truth in the post-photographic era.* Cambridge, MA: MIT Press.

Penny, S. (1995). *Critical issues in electronic media,* New York: SUNY Press.

Willams, R. (1975). *Television: Technology and cultural form,* New York: Schocken Books.

eBooks in the Cloud: Desirable Features and Current Challenges for a Cloud-based Academic eBook Infrastructure

Wei Shen

GESIS – Leibniz Institute for the Social Sciences, Lennéstr. 30, 53113 Bonn, Germany.
Email: wei.shen@gesis.org.

Ute Koch

GESIS – Leibniz Institute for the Social Sciences, Schiffbauerdamm 19, 10117 Berlin, Germany.
Email: ute.koch@gesis.org.

Abstract: *With the rapid development of the mobile technology, the use of multifarious mobile devices, such as tablet PCs, eReader devices and mobile phones for electronic reading, has become an important part of everyday life. In December 2010, Google launched its Google eBookstore with more than 3 million e-books. The most significant feature of the new Google application is that it stores the book content in the cloud, which enables the reader to access the book content at any time on any mobile device in a seamless manner. A cloud-based eBook infrastructure has the characteristics of providing a vast amount of publication information, storing and processing the data in the cloud and displaying the results on demand on diverse desktops or mobile services. Taking into account the characteristics of a cloud-based eBook infrastructure, the paper examines the current eBook distribution lifecycle consisting of eBook publishing, cloud computing and mobile reading technologies. Potential features necessary for a cloud-based academic eBook infrastructure which will better support publishing, searching and reading on the web, as well as promote communication in the scientific community, are outlined.*

Keywords: *Cloud-based eBook infrastructure; semantic publishing; mobile devices; mobile reading; electronic books; collaboration; information search.*

Introduction

Neither mobile technologies nor electronic book publishing is a newly emerging field, but with the success of Amazon's Kindle and the rapid growth of Amazon eBook store, they have received increasing attention in recent years and are becoming the craze fields of innovations.

There is a wide variety of dedicated eBook readers on the market today. EBooks can be read on various portable mobile devices, such as mobile phones, PDAs, tablet PCs. The rapid development of specialized software applications for mobile services (iBook, Stanza and Good Reader) provide new reading possibilities.

In December 2010, the search engine giant, Google, entered the eBook sales world and launched Google eBooks offering more than three million titles. One important feature of the new Google application is the ability to store the content of a book in the Google-Cloud. It enables the reader to access any eBook content "seamlessly at any time on any device" (Google books overview, 2010). The current world's largest eBook online store Amazon unveiled an expanded version of *Kindle for the Web* early in 2011 only one month after Google launched its eBook store.

The idea of a cloud-based model for an eBook system is not new. Safari Books (Version 6.0) released in October 2009 attempted to improve interactivity (e.g. inline notes) and collaboration (categorization sharing) with the basic concept "everything is always in sync because your library is in the cloud" (O'Reilly, 2009).

The Internet Archive's BookServer project (www.archive.org/bookserver) utilizes the cloud concept as well. It allows readers, booksellers, authors, libraries and publishers to access and distribute eBooks available in its distributed system through any devices they have.

The emergence of cloud-based eBook libraries is certain to provide new research opportunities, influence digital publishing and enhance the acceptance of eBooks and other reading devices by readers.

Although eBooks are noted as "the obvious next step to ring a full line-up of web-based basic library resources" (Dillon, 2001), many people remain reluctant to read digital documents from screens (O'Hara & Sellen, 1997). In particular, academic use of eBooks encounters great resistance. Michael Gorman, former president of the American Library Association, believed that "massive databases of digitized whole books, especially scholarly books, are expensive exercises in futility" (Gorman, 2004). Recently, eBooks are being presented in a much more user-friendly manner on computer screens, dedicated eReaders and mobile devices (e.g. free open eBook standard format by IDPF in the digital publishing world and the introduction of E Ink and LCD Display technology for mobile reading devices).

However, the "paper-like" reading experience or effect alone is not enough to convince the conservative book readers to accept eBooks. Instead of being the surrogate for print books, eBooks should take advantage of the dynamic nature of being digital text and provide innovative functionalities which supersede traditional paper books.

Definition and Purpose

In this paper, important and desirable features for cloud-based academic eBook infrastructure are proposed. Taking into account the perspectives of both the eBook publishers and the end users, we have examined current eBook publishing, mobile technologies on the one hand and the search and social reading behavior in scientific communities on the other. The proposed features are considered essential and necessary features in a cloud-based academic eBook infrastructure.

We adopt the definition of cloud in the white paper "A Berkeley View of Cloud Computing", namely cloud is used to define hardware and software in the datacenter, which provides applications as services referred to as Software as a Service (SaaS) over the Internet (Armbrust *et al*., 2009). With this definition of cloud, it follows that a cloud-based model of eBook infrastructure allows storing digital content in the cloud and providing end users with services for discovering, selecting and accessing the eBook they need. Furthermore, it presents end users with the possibility of reading eBooks from anywhere on any device.

Due to the vast number of eBooks in the cloud, end users require and demand that searching and discovering of eBooks in a cloud-based infrastructure be intelligent and precise. Given that nowadays, multiple devices with internet connection (PC, Laptop, Tablet PC, mobile/smart phone, etc.) are increasingly used by end users to access the eBooks on the web, it is a natural step to move eBook relevant services and applications to the cloud. Such a move facilitates communication and cooperation among end users, most notably among the users in the academic community. With eBook relevant services and applications located in the cloud environment, researchers are able to communicate and collaborate with one another at anytime and anywhere and can be benefited by using these services in the cloud during their whole research work process.

The emergence of cloud-based eBook infrastructure has altered the publishing process, from traditional printing to internet digital publishing. Recent studies by Outsell's Gilbane Group emphasize the following five disruptive technologies as crucial for publishing industries: cloud computing, mobile computing, business intelligence, semantic technology and Enterprise 2.0. They expect that these five technologies "will form the bedrock for the next generation of content selling, provisioning, and monetization" (Guenette, Trippe, & Golden, 2010). In addition to reducing cost in the process of adopting cloud-based technologies, publishers also face the challenge of understanding the shifts from the "digitized world" to a "community-centric web world".

In our view, improvement of the mechanisms for digital content discovery is the key to "eBooks in the cloud". Innovative technologies, such as collaborative recommendation and various semantic-based

approaches discussed below, provide both the end users and eBook publishers with possible means of achieving a desired level of efficiency in content discovery. However, the cost efficiency and implementation of such features will need to be evaluated by eBook publishers and eBook vendors.

Using Cloud-based Infrastructure to Select and Identify the eBook Needed

While cloud computing technology facilitates access to a vast number of electronic books on the web, the task of providing a seamless reading environment with good search experience has been of great interest to the researchers.

Content Searching Inside the Book

Full text search is the most used strategy for searching in a book in the electronic environment. But if the reader types a synonym or misspells the search term, the search result is not what s/he is looking for. A semantic search service in the cloud should be the appropriate tool to simplify the search process for great convenience. With the current semantic technology, such semantic search services can operate on the input search term in the eBook cloud environment, predict the term the user is searching for in the book with artificial intelligence and return the searching results to the reader's dedicated reading device properly.

Academic readers often search for footnotes and citations for their research goals. Sometimes a URL link is addressed by the author in the reference list, if the full text paper or electronic book is available on the web. Otherwise, the reader must search the literature via a web search engine, such as Google Scholar, or go to the library to find the literature s/he needs. On account of this, hyperlinking the reference lists of the book, which provides the reader with further information on the references, for example the functionality to indicate in the reference list of a book itself whether the full text of the referenced literature is available in the cloud library, is desirable.

Searching among the eBook Collections in the Cloud

The substantial numbers of eBook collections in the cloud and the capability of consistent access to the content across variable mobile devices, enable the user access to effortlessly and rapidly distributed eBooks on the web, conversely making search among the book collections in the cloud more challenging. The semantic web advocates a revolution that presents the publication information in a meaningful way and facilitates efficient data aggregation, and the capability of consistent access to the content across variable mobile devices accelerates automated information discovery. Therefore approaches of "semantic-aware search" are needed to bridge the gap between the heterogeneous publication data providers over the internet and to improve the readers' search experience.

Ontology is machine processable meta information which describes data semantics (Fensel, 2003). The task of developing a general eBook-cloud ontology model is to conceptualize the eBook-cloud domain knowledge using Web Ontology Language (WOL) (W3C, 2004), which is composed of a set of modular ontologies with relevant domain concepts, and relationship between the concepts. The goal of a general eBook-cloud ontology model is to "enable semantic interoperability between resources, services, databases, and devices via inter-related knowledge structures" (Cheung, Drennan, & Hunter, 2008).

Impact of Semantic Web Publishing

Shotton (2009) stated that semantic publishing aims to enrich the semantic tenor of a publication, enable the automatization of information discovery, compute the semantic relevance to other publications, aggregate the heterogeneous data sources, and promote the interactivity of data within the publication.

As with eJournals[1] STM publishers are vanguards in implementation of new ways in publishing eBooks and exploring new features that were mostly launched first in textbooks within eLearning environments.

If publishers in recent years thought about "writing once, publishing in multiple formats", they should now contemplate making the content readable not only for humans but also for machines. If this can be solved, content will not only be significantly improved but also easily found. In addition, it will enable a broader view on the context of research output, as well as provide substantial benefits to academic communication, particularly when social networks become an integrated part of the eBook cloud.

Utilizing semantic web technologies like RDF (W3C, 2004), ontologies, taxonomies, approaches are already being made to achieve the goal of semantic publishing. The PAUX Company has developed an innovative content management system in which the basic content component is semantic micro-content, linked to each other not by hyperlinks but by "PAUX-Links". These PAUX-Links can be rated and extended to enhance the semantic meaning of the contents, which enables semantic search and filtering.

The cloud-based concept of eBook systems has the potential and points the way: "we need servers which expose existing data systems using the RDF standards, documents 'marked up' and ontologies" so enhanced publications can interact in a semantic way (Vernooy-Gerritsen 2009).

Establishing Multiple Recommendation Services for eBooks in the Cloud

Recommendation tools are often seen as a useful supplement mechanism to information discovery. Wang (1998) defined a recommender system as "a system that has as its main task, choosing certain objects that meet the requirements of users, where each of these objects are stored in a computer system and characterized by a set of attributes".

Cloud-based eBook infrastructure provides readers with a distributed framework not only for sharing and synchronizing over mobile reading devices, but also for opening up an opportunity for deducing and recommending the valuable eBook information that meets the needs of the reader actively.

The following are some proposed recommendation services for an academic eBook cloud-based infrastructure.

A content-based recommendation service should be a useful tool for mobile readers to discover further eBook items that they would be interested in reading. According to content-based filtering algorithms, this service retrieves and analyzes the metadata of the item individuals buy or read and recommends other eBook items with similar content. Typical examples are functionalities such as the "see more related books" button or "new arrivals". The reader can also create a personal profile that instructs the cloud infrastructure to operate his preference inquiry and make recommendations to his end mobile device.

Collaborative recommendation service is "based on collecting and analyzing a large amount of information on users' behavior, activity or preferences and predicting what users will like based on their similarity to other users" (TVGENIUS, 2010). A cloud-based eBook infrastructure is capable of collecting the interactive user information implicitly from web-connected handheld mobile devices or multiple desktops the reader uses. With the help of these data, the collaborative recommendation service can compare and calculate the different interaction attributes such as watch list of the user, previous reading history, the frequency of an eBook access of the reader as well as reading time of an eBook, and finally recommend eBooks stored in the cloud to the user's device.

For academics, recommendations from colleagues or researchers with the same research interest play an important role in the eBook finding process. Social recommendation service is widely used in the social network portals such as Facebook. The eBook cloud-based infrastructure can incorporate such

[1] The "Article of the Future" project from Elsevier and Cell Press, the semantically enhanced version of an original PLoS NTD paper by Reis et al. (2008), undertaken by Shotton, Portwin, Klyne, & Miles (2009), or the Semantic Biochemical Journal are just some of the works worth mentioning here.

a social recommendation service that allows the reader to evaluate the eBooks s/he read and recommend them to the community.

Using Cloud-based eBook Infrastructure to Facilitate Collaboration and Communication among Academic Communities

Book reading in a group is a social activity and requires more "in-use features" (Vasileiou & Rowley, 2008) such as annotations, highlights and bookmarking. For academics who have the same research goals, social reading activity is common and has been proved to exchange research ideas, contributions and research results efficiently. As cloud computing technology offers institutional, corporate and individual users the ability to share storage space and application resources from a central, off-site location accessible from the internet (Carr, 2009), meanwhile the online activity of academic staff and students is increasingly distributed across multiple devices; this centralized facility can promote simultaneous collaboration and allow the researchers independence from their physical location, and the reading devices they use work on shared digital text together. Current applications of cloud-based eBook infrastructure by Google eBook do not support this kind of user collaboration activities in the Google environment; the user can neither bookmark the location in the Google eBook nor highlight the text .

One of the challenges is how to handle the shared references (Marshall, 2009) in eBooks on different reading devices. The traditional page number is apparently still not suitable for shared eBook referencing. Different eBook reading devices transfer the same eBook with different page numbering schemes, and the page number always changes depending on the zoom level. Thus, page numbers are not a consistent way of referencing in mobile devices. However, the eReader producers have already provided possible solutions. For instance, the Kindle eReader adopts document location to identify the specific location within the eBook. Google Reader Software offers another page referencing model, which intelligently displays the actual print page number of the physical book. In the case of a changed zoom level, a page number indicates the exact page number of the print page, such as "page 2-3". However, since such resolutions are all device-specific, interchanges with diverse mobile services are still obviated. In order to allow collaboration among the common reading practices mentioned above in the community, dynamic eBooks should be published with semantics encoded. The basic idea is similar to the "PAUX"-concept: fractionizing the eBook fragments to a fine granularity level. The fragment can be a paragraph, a sentence, or a word, linking these fragments semantically. Furthermore, hard/software of eReader devices or other mobile eBook reading devices should be improved and developed continuously for a better customization possibility in dynamic eBooks.

We have further observed that the process of actively sharing knowledge among academics and students in a social reading group is extremely goal-oriented. It often aims to allow team-work to achieve a specific goal, such as a research paper or a term paper. There are a number of collaboration applications developed for such a purpose. For instance, the application Google Docs enables users to share and work on a document in the Google-Cloud. However, handheld eBook reading devices are designed to enable the user to read and provide little support for the writing activity. This is also the case in Apple iPad, whose virtual touch keys continue to present a certain level of inconvenience, even in writing a short e-mail. An easy-to-use writing functionality is necessary in mobile reading devices.

Despite the hardware problem of writing, the design of collaboration tools, so-called "eBook publishing tools in the cloud", is also very challenging. Specific Content Management Services should be composed to organize and manage the consistency and composition of the content units with varied media type and data size, in order to make them "look and feel like a book", and another important feature for such a collaborative publishing application is certainly the automated apposition of semantic meaning for an enhanced eBook. Furthermore, a reviewing process for quality control is also a necessary application for the Cloud Content Management Service.

Personalization

The prerequisite to enjoying "seamless reading" of Google eBooks is that the reader own a Google user account. Personalization in the cloud environment is a general trend. It is a requisite factor for

many add-value features such as the personalized recommendation service mentioned above. Since storing the personal user profile in the cloud has the advantage of data persistence and consistence across the multiple devices, the personalization feature is highly desirable. Nevertheless personalization also brings challenges in user security. Taking an example of the famous eBook seller Amazon, Amazon remotely deleted eBooks from the customers' Kindles without any prior notification (Stone, 2009). It is imperative that any cloud service providers not only protect the sensitive personal data stored in the cloud, but also obey the privacy requirements of the customer.

Conclusions

Google CEO Eric Schmidt (2007) stated in an interview that "It's pretty clear that there's an architectural shift going on. These occur every 10 or 20 years. The previous architecture was a proprietary network with PC clients called client-server computing. With this new architecture you're always online; every device can see every application; and the applications are stored in the cloud." Cloud computing was one of the technology trends in 2010; it is also expected to be the case in 2011.

When Tom Mulally, founder of Numagic, a consulting company specializing in new media and emerging technology projects, presented a concept for the next generation of content production process, he stated that

> first the cloud would give us instant connectivity to whatever we need, when we need it, with transactions handled automatically. Then semantic search and related technologies would be combined to provide decision support, analytics, reasoning. . .essentially an omniscient intelligent-assistant running in the background. And the multi-sensor capabilities of mobile smart devices (iPhone, Android, iPad, etc.) connected via high speed wireless creates a sort of positive feedback loop, keeping all stakeholders, including the end user/client as active participants throughout the process. (Mulally, 2010)

Cloud computing and mobile technology continue to influence the development of eBook reading. They have the potential of revolutionalizing the traditional publishing industry. We consider it extremely important that approaches be taken to strengthen and improve the technologies in these two areas.

There are a variety of impacts on electronic publishing which we did not consider in detail here. For example James McQuivey from Forrester Research argued "the ultimate effect of Google eBooks. . . will be the creation of an ad-supported publishing model" (McQuivey, 2010). Another change discussed by Hugh McGuire on the O'Reilly radar points to the arbitrary distinction between "the internet" and "books" which he expects will disappear in the near future (McGuire, 2010). Nevertheless, the features and challenges summarized in the paper provide a step forward into an intelligent cloud-based academic eBook infrastructure with the capability of effective publishing, disseminating and sharing electronic resources on the web. These issues continue to be the focus of our work. With the initial findings, we hope to open up new avenues of research in this arena.

References

Armbrust, M., et al. (2009). Above the clouds: A Berkeley view of cloud computing. Technical Report EECS-2009-28, EECS Department, University of California, Berkeley. Retrieved May 2, 2011 from http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf

Carr, J. (2009). Box up your files and move 'em to the cloud'. *Information Today*, *26*, 20.

Cheung, K., Drennan, J., & Hunter, J. (2008). Towards an ontology for data-driven discovery of new materials. In D.L. McGuinness, P. Fox & B. Brodaric (Eds.) *Papers from the AAAI Spring Symposium, March 26-28, 2008, Stanford University, CA.* (Technical Report SS-08-05). Menlo Park, CA: AAAI Press. Retrieved May 29, 2011 from http://www.aaai.org/Papers/Symposia/Spring/2008/SS-08-05/SS08-05-003.pdf.

Dillon, D. (2001). E-books: the University of Texas experience, part 1. *Library Hi Tech*, *19*(2), 113-125.

Fensel, D. (2003). Ontologies: A silver bullet for knowledge management and electronic commerce. Retrieved May 2, 2011 from http://software.ucv.ro/~cbadica/didactic/sbc/documente/silverbullet.pdf

Google books overview. (2010). Retrieved May 2, 2011 from http://www.books.google.com/help/ebooks/overview.html

Gorman, M. (2004). Google and God's Mind: The problem is, information isn't knowledge. Retrieved May 2, 2011 from http://serials.infomotions.com/colldv-l/archive/2005/att-0054/M-Gorman___Google_and_God_s_Mind.pdf

Guenette, D.R., Trippe B., & Golden K. (2010). A blueprint for book publishing transformation: Seven essential processes to re-invent publishing. (Research report of Outsell's Gilbane Group). Retrieved May 2, 2011 from http://www.bisg.org/docs/gilbane_blueprint.pdf

Marshall, C.C. (2009). *Reading and writing the electronic book.* Morgan & Claypool Publishers.

McGuire, H. (2010). The line between book and Internet will disappear. The inevitability of truly connected books and why publishers need APIs. Posted on 10 September 2010. Retrieved January 23, 2011 from http://radar.oreilly.com/2010/09/beyond-ebooks-publisher-as-api.html

McQuivey, J. (2010). Google eBooks paves the way for ad-supported publishing. Posted on 06 September 2010, Retrieved January 23, 2011 from http://blogs.forrester.com/james_mcquivey/10-12-06-google_ebooks_paves_the_way_for_ad_supported_publishing

Mulally, T. (2010). Key emerging technology trends for media professionals. Retrieved January 22, 2011 from http://emerging-technologies-information08.blogspot.com

O'Hara, K. & Sellen, A. (1997). A comparison of reading paper and on-line documents. In *CHI '97: Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 335-342). New York, NY. ACM. Retrieved May 29, 2011 from http://old.sigchi.org/chi97/proceedings/paper/koh.htm.

O'Reilly, T. (2009). Safari Books Online 6.0: a cloud library as an alternate model for ebooks. Retrieved December 20, 2010 from http://radar.oreilly.com/2009/10/safari-books-online-60-a-cloud.html

Reis, R.B., Ribeiro, G.S., Felzemburgh, R.D.M., Santana, F.S., Mohr, S., et al. (2008). Impact of environment and social gradient on Leptospira infection in urban slums. *PLoS Neglected Tropical Diseases*, 2: e228. Retrieved May 2, 2011 from http://dx.doi.org/10.1371/journal.pntd.0000228 .

Schmidt, E. (2007). Text of Wired's interview with Google CEO Eric Schmidt. Retrieved May 2, 2011 from http://www.wired.com/techbiz/people/news/2007/04/mag_schmidt_trans

Shotton, D. (2009). Semantic publishing: the coming revolution in scientific journal publishing. *Learned Publishing, 22,* 85-94.

Shotton D, Portwin, K., Klyne, G., & Miles, A. (2009). Adventures in semantic publishing: Exemplar semantic enhancements of a research article. *PLoS Computational Biology*, 5(4): e1000361. doi:10.1371/journal.pcbi.1000361.

Stone, B. (2009). Amazon erases Orwell books from Kindle. *The New York Times*. Retrieved May 2, 2011 from http://www.nytimes.com/2009/07/18/technology/companies/18amazon.html?_r=1

TVGENIUS. (2010). TV recommendations using relevance maps. Retrieved May 2, 2011 from http://www.tvgenius.net/resources/white-papers/an-integrated-approach-to-tv-recommendations/

Vasileiou, M. & Rowley, J. (2008). Progressing the definition of "e-book". *Library Hi Tech*, 26(3), 355-368.

Vernooy-Gerritsen, M. (2009). *Enhanced publications*. Amsterdam: University Press.

W3C (2004). OWL Web Ontology Language overview. Retrieved May 2, 2011 from http://www.w3.org/TR/owl-features/

Wang, P. (1998). Why recommendation is special? In *Papers from the 1998 Workshop on Recommender Systems, part of the 15th National Conference on Artificial Intelligence (AAAI 1998)* (pp. 111-113), Madison, Wisconsin: EUA.

If You Can't Retrieve it, Does it Exist? Accessibility of LIS Journals on the Internet

Tove Faber Frandsen

University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark.
Email: tofr@litcul.sdu.dk

Jeppe Nicolaisen

Royal School of Library and Information Science, Birketinget 6, DK-2300 Copenhagen S., Denmark.
Email: jni@iva.dk

Abstract: *Public access to the World Wide Web became widespread in the late 1990s and early 2000s, and today documents are frequently published on the Internet. Open access (OA) to the scientific literature has been found to be increasing as more and more producers and publishers of scientific literature make their publications available free of charge on the Internet. The paper puts forward that it may be argued that only scholarly Internet documents that are retrievable through the search engine Google Scholar (GS) can be said to exist. The degree of coverage of GS is thus an important issue. The paper reports the results of a study of 159 journals in the field of Library and Information Science and their degree of coverage in GS. Journals publishing many issues a year are not found to be more retrievable than journals with fewer issues. Non-English and OA journals tend to have a lower degree of retrievability. The tendency is found to be even stronger for journals that are both OA and non-English. OA and non-English journals are very heterogeneous groups and the variation in their degree of retrievability is found to be much higher than in the case of traditional, toll-access journals, which resemble each other more in relation to retrievability.*

Keywords: *Google Scholar; retrievability; open access.*

Introduction

There is a famous riddle in philosophy that goes something like this:

> *If a tree falls in a forest and no one is around to hear it, does it make a sound?*

The riddle can be traced back at least to George Berkeley (1685-1753) and his work "A Treatise Concerning the Principles of Human Knowledge". Raising questions, as it does, about observation and reality, the riddle has occupied many bright thinkers ever since. The basis of the riddle is, of course, whether something can exist without being perceived. We shall not attempt to solve the riddle here, but only mention it in passing as it resembles a similar riddle we have discovered in our own field of Library and Information Science (LIS):

> *If a document is published but nobody can find it, does it exist?*

Imagine, for instance, a misplaced book in a library collection. The classification code in the library catalog tells a user where to find the book, but the book is not there. It was mistakenly placed at the wrong bookshelf by the last user. Does the book exist then? Just like the riddle about the tree falling in the forest, this riddle concerns the question whether something can exist without being perceived. It is a metaphysical question, and the answer depends at least in part on whether one believes that an object is distinct from its properties or whether an object is merely its sense data.

Taking the question about existence to a more concrete level, a number of studies have shown that scholars and scientists tend to adopt easily accessible information resources to a larger extent than less easily accessible information resources. One example is a study of neuroscientists performed by Vibert et al. (2007). This study suggests that when scientists can access efficient and exhaustive BDI resources online, those resources quickly become their preferred way of getting work-related

information. This corresponds to the findings of Tenopir et al. (2005) who found that astronomers tend to select access means that are convenient whether in print, electronic, or both. Thus, the prediction put forward by Odlyzko (2002) that scholars, publishers and librarians will have to make even greater efforts to make their material easily accessible is as true today as it was almost ten years ago. In recent years there has been a major change in the ways scientific communication (both formal and informal) is disseminated by electronic means. At the same time, the information seeking behavior of scientists have changed, especially when it comes to the use of search engines and digitized resources apart from traditional journals (Meyer & Schroeder, 2009). The electronic accessibility of journals implies, among other things, that scientists now make fewer library visits. Web browsing and table of contents e-mail alerts are gradually replacing physical browsing, and searching has been found to be a popular option for keeping up to date with new developments (Olle & Borrego, 2010). However, like the misplaced book in the library collection, one may ask whether this open access (OA) literature really exists. Is 'putting something on the Internet' equal to an existence claim? We do not believe so. The Internet is humongous. Consequently, putting something on the Internet is in itself no existence claim. It's like the misplaced book in the library collection. There is no guarantee that the publication will ever be found again. Without going deeper into the metaphysical question about existence, we believe that a first requirement for the existence of an Internet document is that it is retrievable. This, of course, raises a question about what it means to be retrievable on the Internet. All documents on the Internet have a unique URL. In principle, all documents on the Internet are thus retrievable – one only needs to know the correct URL. Again, this compares to the misplaced book in the library collection. One could find it if one knew where to look. Thus, a unique URL is not enough. Limiting the discussion to scientific documents, we will argue that a document is retrievable if it can be found by searching a proper search engine. What, then, is the proper search engine for scientific documents on the Internet? We believe it is Google Scholar (GS) (scholar.google.com). This belief is based in part on its presumed preference among scientists (Jamali & Asadi, 2010) and in part on reports stating that GS ensures a higher precision in the search results compared to the comparable search engines due to its more narrow definition of scholarly literature:

> Similar in approach, but broader and less specific in scope than Google Scholar, the scientific search engine Scirus (www.scirus.com) searches, according to information they provide, approximately 300 million science-specific web pages. In addition to scientific documents from Elsevier (ScienceDirect server, see www.sciencedirect.com/) freely accessible documents are provided, many from public web servers at academic institutions. Among these are, for example, documents placed by students that do not fulfill scientific criteria such as peer review, which often lead to their exclusion in searches. In our experience there is more than a negligible fraction of records from Google Scholar non-academic web spaces in the Scirus index. Scirus' coverage of purely scientific sources in addition to Elsevier's ScienceDirect full-text collection is low by comparison (compare the selection of hosts in the Scirus advanced search interface, http://scirus.com/srsapp/advanced/). What Scirus declares as the "rest of the scientific web" is too general, non-specifically filtered and makes up the majority of hits in any query. (Mayr & Walter, 2007, pp. 815-816)

Consequently, if one wants to find scientific documents on the Internet, GS is currently the place to go and search.[1] But how well does GS perform? To what extent does it cover the scientific literature on the Internet?

A study by Chen (2010) finds that GS covers as much as 98 to 100 percent of scholarly journals from both publicly accessible Web contents and from subscription-based databases that GS partners with. Lewandowski (2010) measures the coverage of GS for Library and Information Science (LIS) journal literature as identified by a list of core LIS journals from a study by Schlögl and Petschnig (2005). He checked every article from 35 major LIS journals from the years 2004 to 2006 for availability in GS. The information on type of availability (whether a certain article was available as a PDF for a fee, as a free PDF or as a preprint) was analyzed and divided by the type of publisher. However, the study does not include open access as a specific characteristic of the journals in the analysis. Lewandowski (2010, p. 257) states:

---

[1] We acknowledge that there are also limitations and problems with GS (see e.g. Jacso (2009; 2010) and http://en.wikipedia.org/wiki/Google_Scholar#Limitations_and_criticism), but to our knowledge no other search engine currently matches up. In some fields Google Scholar performs better in terms of both recall and precision than most of the subscription databases when submitting simple keyword queries (Walters, 2009) whereas in others Google Scholar matches the results of a subscription database in terms of recall but not precision (Anders & Evans, 2010; Freeman et al., 2009).

> The numbers for the preprints are somewhat disappointing as there are high hopes for open access (OA) and the willingness of authors to make their work available through OA. Particularly for the LIS profession with its many OA promoters, the numbers seem to be very low.

The conclusion by Lewandowski is supported by Way (2010) who concludes that "the archiving of articles is not a regular practice in the field" (p. 302). Way (2010) measures the open access availability of LIS research in GS by searching for articles from 20 top LIS journals. The study does not focus on open access journals but concludes that GS is an effective search tool for retrieving LIS articles.

Open access can also be in the form of open access journals and, consequently, we would like to address the issue about GS coverage in the present study by including a substantial number of open access journals. The paper is, of course, a coverage study of an Internet search engine, but as we have tried to argue, our topic also concerns deeper philosophical questions.

Methods

With this study we want to investigate whether the open access LIS journal literature is retrievable in GS regardless of the form of publication indexed (e.g. preprint or publisher version). First of all we need to draw a sample of LIS literature on the basis of some formal characteristics. To generate our dataset we used the list of LIS journals available on Journal Citation Reports (JCR) and Directory of Open Access Journals (DOAJ). The former database list includes 66 journals and the latter 117. Excluding duplicates, our dataset for the current study consisted of 177 journals (see Table 1 for the journal list). For the present study all document types are included. The number of publications published in 2009 was checked on the journals' web sites. No distinction was made between various publication types. Obviously some publication types report research to a greater extent than others. Only publications such as call for papers or author instructions were excluded from the study. In the discussion and conclusion section we will return to how future work could strengthen the data set.

Now we need to determine the variables to be included in the study.

The dependent variable is number of retrievable documents in GS. An operationalization of the concepts of "visibility", "retrievable" or "indexed by GS" is necessary. A publication can hardly be said to be visible and retrievable, if it only appears in GS with bibliographic information in a bibliography. The number of subject access points is relatively impeded and we will not gain access to the full publication, an abstract or subject headings. In this study a document is determined to be retrievable from GS if the document is indexed with subject access points beyond the bibliographic reference, i.e. abstract, full text or subject headings. In some cases toll access prevents us from gaining access to the full document although it is indexed in GS. Each journal was checked in GS for the publication year of 2009. Noruzi (2005) summarizes the search techniques available through GS. Peter Jacso (2005; 2008) argues that the size of the database has increased substantially from 2005 to 2008 and GS primarily lacks satisfactory levels of correct assignment of journal names and authors which causes difficulties in the retrieval of articles. However, random examination indicates that the problem is atypical in this case. In this case the queries submitted were based on journal name and publication year using the advanced scholar search formular.

Furthermore, we need to determine the appropriate independent variables. Forms of open access could be an independent variable. Lewandowski (2010) finds that the availability of preprint versions of articles is somewhat disappointing as only a low number of articles is available in this form. The author finds that it is not caused by the indexing by GS but it is rather a result of the lack of self-archived publications by LIS authors. This leads the author to conclude that "while many researchers and practitioners in the LIS field are advocates of open access, when it comes to advancing open access through depositing preprints, their support is limited" (Lewandowski, 2010, p. 260). Consequently, we choose not to separate availability in forms. We focus on whether the articles are indexed (regardless of whether there is open access to a full text version).

Secondly, language needs to be considered as an independent variable. Lewandowski (2007, 2010) finds that a much higher percentage of English-language LIS articles is available in GS. However, the non-English journals are typically published mainly by small publishers and, for some, no online version is available, therefore there is a lower coverage rate. The author concludes: "The results do

not indicate a language or country bias towards English-language articles" (Lewandowski, 2010, p. 260). We would like to extend the study of non-English journals to other languages than German and to non-English journals with an online version. Consequently, language is added as a variable in the data set.

The methods of our study followed this procedure: Each journal was entered into the data set with ISSN to uniquely identify the journal, journal name, a dummy variable for language (non-English=1 if journal included publications in other languages than English, otherwise 0), dummy variable for open access (OA=1 for open access, otherwise 0), number of publications in 2009 (according to the journal web sites or alternatively the citation indexes) and number of publications indexed in GS. On the basis of that data we can calculate the share of publications indexed in GS.

Results

First of all, we need to consider a potential correlation between number of publications published and the share of these available through GS. One might argue that small journals (i.e., publishing fewer issues annually) find it difficult to attract attention generally and more specifically by GS whereas large journals (i.e., publishing more issues annually) are generally more visible. Figure 1 illustrates the relation between number of publications and the share available in GS. One outlier has been excluded from the figure as it distorted the picture significantly.

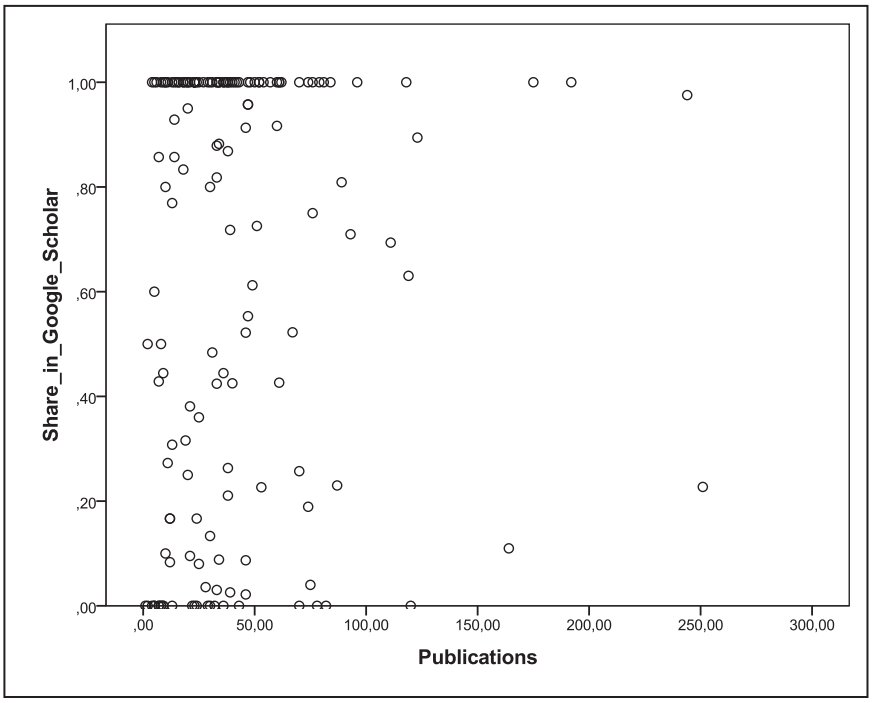

Figure 1. Share in GS and number of issues published annually by the journal

From the figure it is clear that such correlation cannot be found. Statistical analyses confirm the picture depicted in the figure. Linear regressions (with and without the outlier) with share in GS as the dependent variable and publications as the independent variable reveal p-values of the independent variable ranging from .25 and .29. Consequently, we move on with the analyses.

Table 1 provides an overview of the mean and median values. Mean is typically used to describe the central tendency of a set of data that does not have extreme values (outliers). On the other hand, the median is typically used to describe the central tendency of a set of data that does have extreme values as it is not affected as strongly as the mean by outliers. In this case the values range from 0 to 1 and there are no outliers (see Figure 1 and Table 1). Consequently, the mean is the appropriate measure to describe the middle.

Table 1. Overview of results

Share in Google Scholar							
		N	**Mean**	**Median**	**Percentile 05**	**Percentile 25**	**Percentile 75**
Open_access	0	62	.79	1.00	.10	.63	1.00
	1	97	.52	.50	.00	.04	1.00
Non_English	0	108	.68	.88	.00	.40	1.00
	1	51	.51	.26	.00	.00	1.00
OA_Non_English	0	114	.68	.88	.00	.36	1.00
	1	45	.50	.23	.00	.00	1.00

In the following figures we use the mean to analyse differences in coverage. Figure 2 illustrates the visibility in GS of OA journals and traditional toll access journals. The error bars represent the 95% cent confidence interval. We can clearly see that although there is some variance among journals within the two categories of journals, traditional toll access journals have higher coverage rates in GS than OA journals. The former have a mean coverage of 79% cent whereas the latter is 52%.

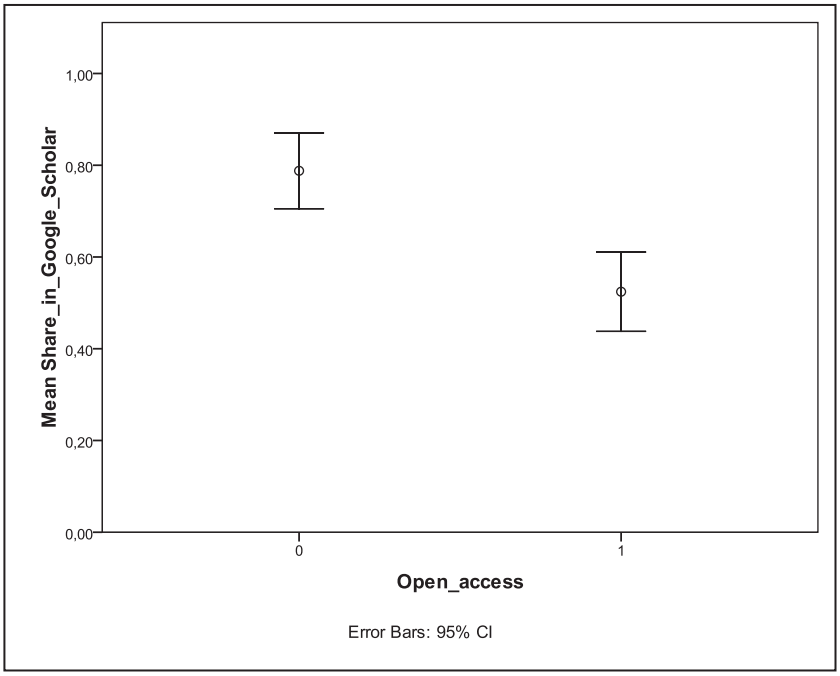

Figure 2. Mean share indexed in GS and open access status. The error bars represent the 95% cent confidence interval

Figure 3 illustrates the visibility in GS of journals that are published exclusively in English and journals that are published in English as well as in other languages. Journals published exclusively in English have higher visibility in GS as their share in GS is 17 percentage points higher than those published in other languages than English. The latter group seems to be a more heterogeneous group as their variance is greater.

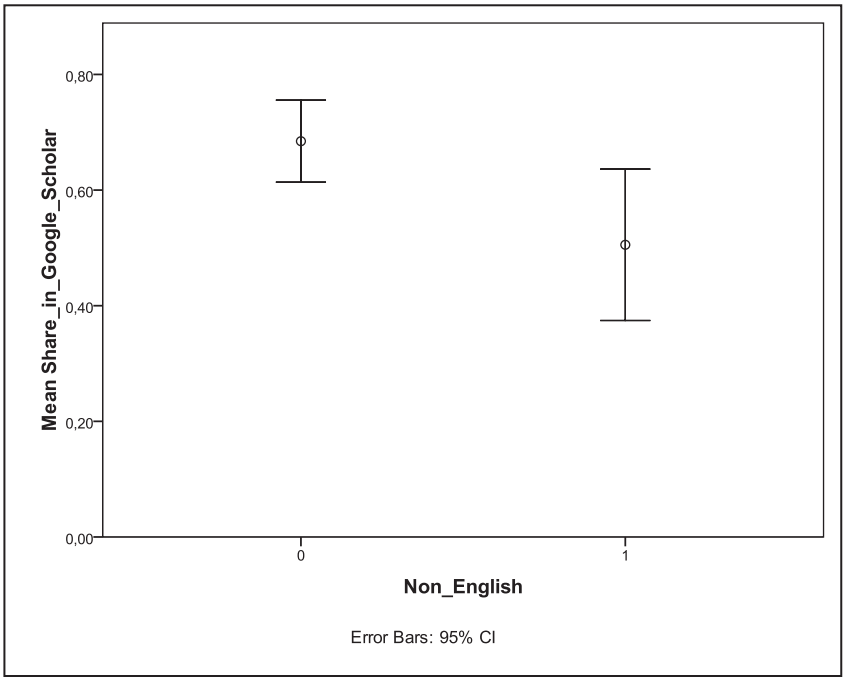

Figure 3. Mean share indexed in GS and language. The error bars represent the 95% cent confidence interval

Figure 4 illustrates the visibility in GS of OA journals that are published in English as well as in other languages compared to toll access journals that are published exclusively in English.

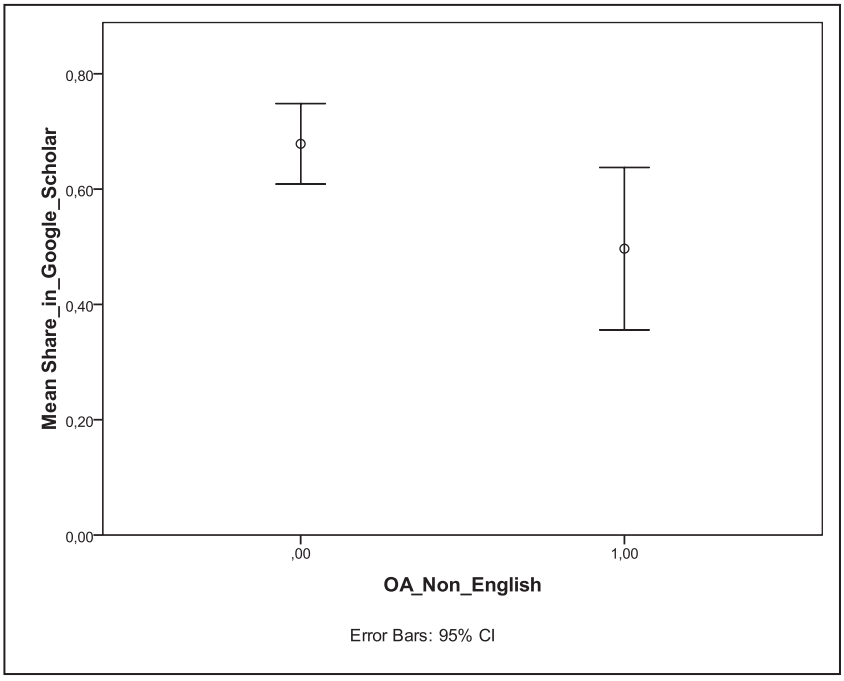

Figure 4. Mean share indexed in GS and OA as well as non-English. The error bars represent the 95% cent confidence interval.

The open access journals including publications not in English are characterised by an even lower share of visibility in GS. The remaining journals of the data set cannot be characterised differently in terms of mean than the non-English journals. Their share is, however, 11 percentage points lower than that of the toll access journals.

Discussion and Conclusion

Before engaging with the discussion of the implications of our results, we will briefly summarize our findings. First of all, large journals (i.e., journals with many issues published every year) are not more retrievable in GS than smaller journals. Non-English and OA journals tend to have a lower degree of retrievability in GS. The tendency is even stronger for journals that are both OA and non-English. In addition, OA and non-English journals are very heterogeneous groups and there is much more variation in their degree of retrievability than in the case of traditional, toll-access journals, which resemble each other more in relation to retrievability.

As we have tried to make it clear throughout the paper, OA is also a question about retrievability, but there need not be a correlation between open access and retrievability. Putting something on the Internet free of charge is generous, but if nobody is able to locate and access it, it will hardly make a difference.

The present work could be extended and strengthened by including more journals in the study. The present study casts light on the relationship between open access and retrievability for open access and journals with enough international impact to be included in JCR. A number of journals with different characteristics than the journals in this study such as toll access journals from developing countries could help depict a richer picture of the relationship between open access and retrievability.

Using GS as our index of retrievability, we have found large differences in the retrievability among OA journals. The responsibility for this lies with the editors. It seems that some OA journals could benefit from a more active approach to having their content indexed by GS. There are technical inclusion guidelines as well as an inclusion request form available at the GS website (http://scholar.google.com/intl/en/scholar/publishers.html#faq1). This study stresses that editors of OA journals should be aware of the importance to comply with these guidelines because in a way: *If you can't retrieve it, it doesn't exist!*

References

Anders, M.E. & Evans, D.P. (2010). Comparison of PubMed and Google Scholar Literature Searches. *Respiratory Care*, 55(5): 578-583.

Chen, X.T. (2010). Google Scholar's dramatic coverage improvement five years after debut. *Serials Review*, 36(4): 221-226.

Freeman, M.K., Lauderdale, S.A., Kendrach, M.G., & Woolley, T.W. (2009). Google Scholar versus PubMed in locating primary literature to answer drug-related questions. *Annals of Pharmacotherapy*, 43(3): 478-484.

Jacso, P. (2005). Google Scholar: The pros and cons. *Online Information Review*, 29(2): 208-14.

Jacso, P. (2008). Google Scholar revisited. *Online Information Review*, 32(1): 102-14.

Jacso, P. (2009). Google Scholar's ghost authors. *Library Journal*, 134(18): 26-27.

Jacso, P. (2010). Metadata mega mess in Google Scholar. *Online Information Review*, 34(1): 175-191.

Jamali, H.R. & Asadi, S. (2010). Google and the scholar: The role of Google in scientists' information-seeking behavior. *Online Information Review*, 34(2): 282-294.

Lewandowski, D. (2007). Nachweis deutschsprachiger bibliotheks- und informationswissenschaftlicher Aufsätze in Google Scholar. *Information Wissenschaft und Praxis*, 58(3): 165-168.

Lewandowski, D. (2010). Google Scholar as a tool for discovering journal articles in library and information science. *Online Information Review*, 34(2): 250-262.

Mayr, P. & Walter, A.-K. (2007). An exploratory study of Google Scholar. *Online Information Review*, 31(6): 814-830.

Meyer, E.T. & Schroeder, R. (2009). The world wide web of research and access to knowledge. *Knowledge Management Research & Practice*, 7(3): 218-233

Noruzi, A. (2005). Google Scholar: The new generation of citation indexes. *Libri*, 55(4): 170-180.

Odlyzko, A. (2002). The rapid evolution of scholarly communication. *Learned Publishing*, 15(1): 7-19.

Olle, C. & Borrego, A. (2010). A qualitative study of the impact of electronic journals on scholarly information behavior. *Library & Information Science Research*, 32(3): 221-228.

Schloegl, C. & Petschnig, W. (2005). Library and information science journals: an editor survey. *Library Collections, Acquisitions, & Technical Services*, 29(1): 4-32.

Tenopir, C., King, D.W., Boyce, P., Grayson, M., & Paulson, K-L. (2005). Relying on electronic journals: Reading patterns of astronomers. *Journal of the American Society for Information Science and Technology*, 56(8): 786-802.

Vibert, N., Rouet, J-F., Ros, C., Ramond, M., & Deshoullieres, B. (2007). The use of online electronic information resources in scientific research: The case of neuroscience. *Library & Information Science Research*, 29(4): 508-532.

Walters, W. H. (2009). Google Scholar search performance: Comparative recall and precision. *Portal-Libraries and the Academy*, 9(1): 5-24.

Way, D. (2010). The open access availability of library and information science literature. *College & Research Libraries*, 71(4): 302-309.

iPhone Mobile Application Design: The Case of Hacettepe University Libraries[1]

Müge Akbulut

Department of Information Management, Hacettepe University, 06800 Beytepe, Ankara, Turkey
Email: mugeakbulut@gmail.com

Abstract: *Thanks to the features of the mobile devices which have been constantly evolving, there has been a diversity of mobile applications. It is highly possible to find a mobile application to carry out any transaction in an application store. Users, whose expectations have been increasing, hope that libraries will be accessible and controllable. Hence, libraries should move towards mobile platforms so that users can easily get access to them. In this study, we first review the criteria that should be paid attention when designing mobile applications for iPhone and then describe the iPhone mobile application that we developed for Hacettepe University Libraries.*

Keywords: *Mobile application design; mobile usability; iPhone application.*

Introduction

Mobile devices, which have become one of the most indispensible parts of our life, allow us to access social networks, to shop and to carry out banking transactions independently from space and equipment. Internet research carried out by Cisco Company reveals that mobile video will be 66% of the world's mobile data traffic by 2014, with a 131% growth between 2009 and 2014 (Cisco, 2009). The Internet Report of Morgan Stanley Research shows that the users who have Internet access through mobile devices will outnumber the users who have Internet access through desktop computers in 2013 when the current growth rate is taken into consideration (Morgan Stanley, 2009).

Today, mobile devices have turned into a universal device for every kind of communication (Rainie and Anderson, 2008). Thanks to the developing technology, the significance of mobile applications in solving such problems as security, bandwidth, location detection and data capacity has increased. This rapid development process caused the users to come up with more demands. It is not enough for them to find what they look for on the network; they also want to be able to find "just-in-time" information without changing their location (Johnson et al., 2010), usually through the use of their mobile devices (Anderson & Wolff, 2010).

At this juncture, institutions providing information services should take innovative steps toward the reorganization of their services and render these services to users in order to satisfy their information needs. For libraries, applications should be devised so that they can be accessible even through mobile media (Dempsey, 2009).

The most difficult part of mobile application design is that content should be given on a small screen. This means that there is no place for something which is useless and has no value for users. In order to achieve this, the requirements, expectations and behaviors of users should be taken into account and the design process should proceed in harmony with users' expectations (Biel, Grill, & Gruhn, 2010).

In this study, the criteria for iPhone mobile application design are elaborated. A mobile application developed for Hacettepe University Libraries (Beytepe Library, Medical Library and the Library of the Ankara State Conservatory) is described (Hacettepe, 2010).

[1] The iPhone mobile application for Hacettepe University Libraries was developed to satisfy the requirements of an undergraduate course (BBY 405) offered by the Department of Information Management of Hacettepe University during the Fall Semester of 2010-2011.

Mobile Media and Libraries

Tomi Ahonen (2008) focuses on advantages and disadvantages of mass media along with their importance in our social life. He defines mobile devices as the genuine mass medium coming after the printed press, recordings (sound recordings, CDs, DVDs, etc.), cinema, radio, television and the Internet. While mobile devices have the features of the other six media, they are different in terms of being personal, always carried, and their "always on" feature.

Smartphones such as iPhone and Android have the feature of sharing audio, text, or image files simultaneously. Along the way, applications developed for mobile devices offer a rich experience to the users (Rayport, 2009). Facebook, the most frequently used social network site, has approximately 250 million users with mobile access. The activity of these users is twice that of users with desktops (Facebook, 2011).

Users having constant access through their mobile devices need more complicated information services and expect libraries to be easily accessible, dynamic and controllable. At this point, libraries should take steps to change their way of rendering information service and to carry their services to the environment of the users (Dempsey, 2009). Users are ready to accept the library services that they are used to. Hence, mobile libraries (m-libraries), the sources of which can be accessed through the use of mobile devices, are one of the wisest solutions.

When library services are ready for mobile use, user accounts, digital reading rooms, electronic books, journals, articles, etc. can be attained during mobility as well (Needham & Ally, 2008).

The iPhone Mobile Application

With the opening of mobile application stores, mobile phones have become a computer platform. Mobile phones are turned into devices with which you can play games, use GPS and watch movies. Users began to shop, carry out their banking transactions and use social networks with the use of mobile applications (Anderson & Wolff, 2010).

If you are a user of Apple App Store, which has 99.4% of the whole market, you will not experience any difficulty finding an application (Foresman, 2009). The App Store has more than 350.000 applications (Apple, 2011). Such applications, developed in such an environment, must be designed in a sophisticated way that will distinguish them from their rivals. Things that should be watched during the design process of iPhone mobile applications might be itemized as: user-friendly interface, target group and a kind of different attitude about its use.[2]

Content

The most difficult part of a mobile application design is that content should be given on a small screen. There is no place for a trivial detail which confuses the users and is useless on a 3,5-inch sized screen. In order to benefit from this small screen, content should be rendered in a simple way which will not puzzle users.

Mobile information architecture is a process of constituting a sketch and structure to mark out the information which mobile sites and applications will have to encompass (Fling, 2009; Kukulska-Hulme, 2007). Applications developed in harmony with this architecture are easier to use in terms of practicality and more successful in terms of serving the purpose of the users.

The size of the content is also significant. Mobile applications are programs with a small size. If they begin to take up more space, they shut down automatically. Therefore, contents like decorative screen images and flashes should be avoided and pictures should be used in a compressed way (Clark, 2010). Applications taking up too much space are also problematic in terms of accessibility. It takes time to install programs since mobile devices connect to the Internet via 3G or Edge speed (Fling, 2009).

---

[2] There are many sources examining the mentioned features (Biel et al., 2010; Fling, 2009; Group, 2009; Nielsen, 2009; Owoh, 2010). Although different categorizations are made under different headings, criteria discussed in these sources largely correspond.

User-Friendly Interface

Accessing the content and using it in an effective way is as important as the content. The most crucial point while rendering the content is the existence of a usable[3] interface which will ensure the delivery of information in the shortest time (Uçak & Çakmak, 2009).

If the use of an application needs a huge time investment or seems to, then its use in the long term is a marginal probability (Krug, 2006). Research shows that users analyze mobile applications that they can download in less than 10-15 seconds. If the interface seems a bit complicated, they try another application which has a similar content (Fling, 2009).

Users benefit from what is rendered to them (some features like menu, content, form, graphics, etc.) in order to attain their aims (Morville & Rosenfeld, 2006). If more emphasis is given to usefulness during the design of such features, users will go through a less annoying process in which they will be more satisfied (Krug, 2006). In a well-done design, a user will know what it is and how it is used by just a glance at one page, and he/she might do that without concentrating on it much. It is a kind of utopia to suppose that users might understand all of the functions of a system when these functions are given just on the interface (Nielsen, 1993). What kind of a reaction will be seen might change from user to user, but intensive pages and background image are a little bit problematic. Thus, simplicity should always be in the foreground.

Another quality criterion about usefulness is the harmony of web pages with mobile browsers. Web pages for mobile devices should be designed by taking into account the small screens of cell phones, smart phones, iPads, and so on. For instance, mobile devices cannot handle web pages where information is presented in multiple columns.

Target Group

The determination of the target group is significant since this plays an important role in the identification of the practicality, and as a result of this, of providing usefulness and service quality. What users need, what their priorities are, how often and when they will use the application, are among the essentialities during the design process.

While designing an application, what has been aimed at by the designer might be different from the need of users. Called "design blindness" by Nielsen (1993), this might be defined as the situation in which the designer supposes that the design of the interface is flawless and excellent. For example, designers suppose that users will browse the pages carefully, and choose the best by analyzing all the choices. But what users choose is the logical option rather than the best one[4] (Klein, 1999).

A Different Kind of Usage Technique

What is important while developing a mobile application is how the content is presented. Some features like navigation, camera, microphone, gravity sensor (accelerometer) which are found in iPhone, might create an advantage during the presentation of the content. For example, when users make use of an application called SoundHound to search a music database, they do not have to type words. If sound is introduced through a microphone, then a search on the database is carried out and information about a song is retrieved.

iPhone Mobile Application for Hacettepe University Libraries

For someone on the move, sources and services of the libraries have become accessible through an application. Also, thanks to some features like barcode scanning and navigation, this process has become more functional.

---

[3] Usability is defined as a level of users' performing their goals in an efficient, comfortable, and acceptable way and is the main factor in measuring the success of system interfaces.

[4] Klein asserts that people (firemen, pilots, etc.) who have time pressure, ambiguous goals, and variable conditions, take decisions which are risky and prefer to choose what they think is the most logical one at first glance instead of focusing on the possible solutions and choosing the best one. Also, he alleges that web users are not looking for the best options, because they are usually in a hurry and think they will not suffer as a consequence of guessing wrong. In Nielsen's practicality tests, the frequent use of the backspace button might be given as proof of this.

We used the mobile libraries report prepared by Cambridge University to identify services to be included in our mobile application (Fig.1) (Needham & Ally, 2008). Authors of the report asked users about library sources and services they want to make use of when they are mobile. Users wanted to get information about the working hours of the library, its place on a map, its contact information, online catalogue and lending information.

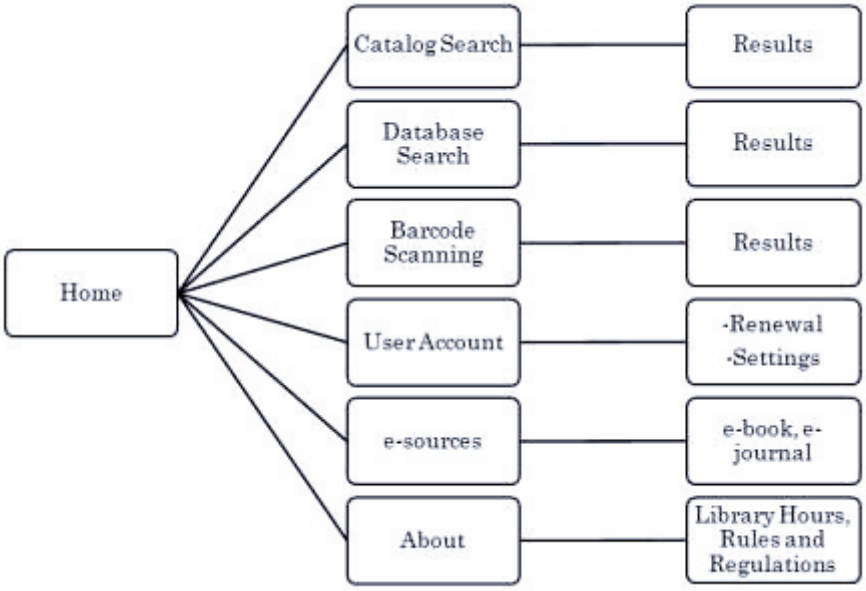

Figure 1. Map of the library application

Barcode Scanning[5]

When the barcode of a book is scanned,[6] it is possible to see information about the nearest place where this book might be found along with a list of opinions about it. In order for this feature to be used in an effective way, an Internet connection is needed (Wi-fi, 3G or Edge). In Figure 2, the barcode of a book titled "Information Architecture for the World Wide Web" is scanned.

Local Searching

In this feature, if the book whose barcode was scanned is found in Hacettepe University Libraries, it is given in the first line. Then, by making use of navigation features, nearest libraries owning the book are given in accordance with their distance (Fig. 3).

[5] Barcode scanning technologies have been applied to the system with the use of RedLaser SDK (http://www.redlaser.com/).
[6] Barcodes can be scanned in the form of UPC-A, UPC-E, EAN-8, EAN-13, QR Code (square code), Code 39, Code 128 and ITF.

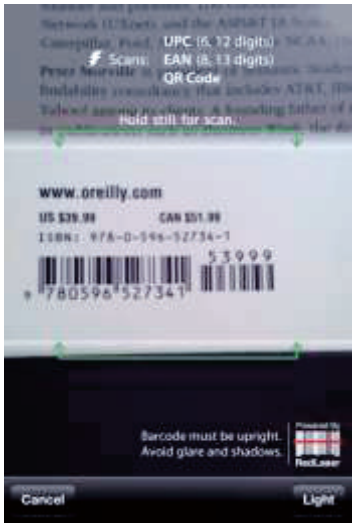

Figure 2. Screen of barcode scanning.

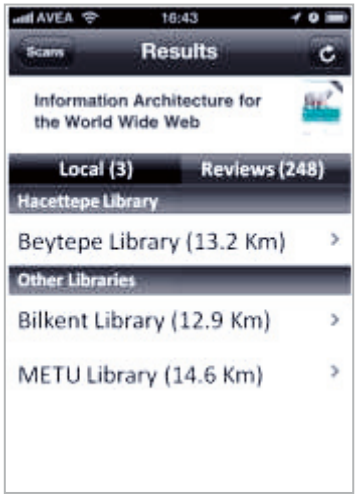

Figure 3. Results of barcode scanning and tab of local search.

When clicked, a connection to the web site of that library is established and relevant information about the book (e.g., call number, where it is located) is rendered (Fig. 4).

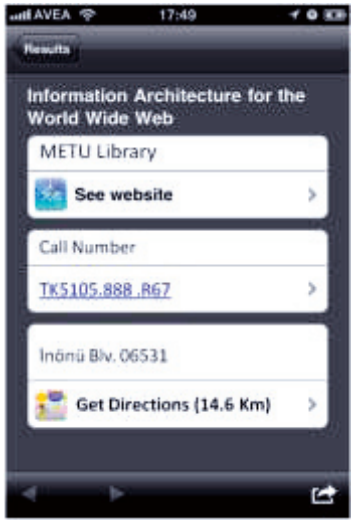

Figure 4. Information for METU Library

When 'Get Direction' tabs are clicked, the directions to get to the library are provided (Fig. 5).

Figure 5. Screen of way finder

Book Reviews

On the screen where the results of barcode scanning are listed (Fig. 3), when the comment tab is clicked, comments about the book along with the number of comments on sites like Amazon, GoogleBooks, WorldCat and LibraryThing are listed[7] (Fig. 6).

Figure 6. Results of barcode scanning, tab of reviews

When the scan button is clicked, the lists of the books whose barcodes have been scanned beforehand are given (Fig. 7). From results, favorite choices might be marked and sent via e-mail.

[7] Barcodes of the books hold ISBN information. Hence, information about the book, cover, comments taken from related sites are formed from ISBN information.

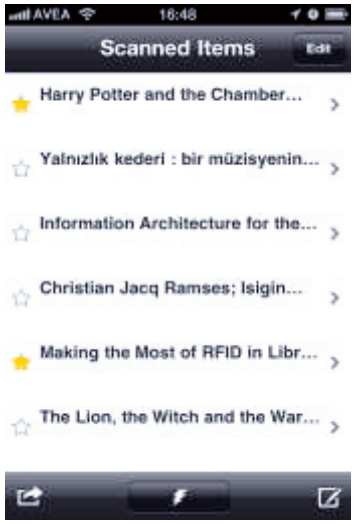

Figure 7. Screen of barcodes scanned beforehand

Conclusion

The Hacettepe University Libraries Mobile Application provides users with an opportunity of accessing library sources through iPad, iPod Touch[8] and iPhone. This application, which was developed in accordance with mobile information architecture and practicality principles, has gained a structure which is useful as an outcome of the features of barcode scanning and mapping the nearest libraries. Users are provided with more diverse information about what they are looking for in comparison with that of libraries (comments available on the Internet about the book, where a particular book might be found elsewhere, etc.). Through this application, some services and sources of the libraries have been moved to a mobile environment.

Acknowledgements

I am grateful to my teachers Prof. Dr. Yaşar Tonta and Dr. Umut Al for their invaluable support and guidance throughout this study.

References

Ahonen, T. (2008). *Mobile as 7th of the mass media: cellphone, cameraphone, iPhone, smartphone*. London: Futuretext.

Anderson. C. & Wolff, M. (2010). The Web is dead. Long live the Internet. *Wired*, September 2010. Retrieved April 2, 2011 from http://www.wired.com/magazine/2010/08/ff_webrip/all/1

Apple. (2011). Over 300.000 ways to make iPhone even better. Retrieved April 2, 2011 http://www.apple.com/iphone/apps-for-iphone/

Biel, B., Grill, T., & Gruhn, V. (2010). Exploring the benefits of the combination of a software architecture analysis and a usability evaluation of a mobile application. *Journal of Systems and Software*, 83(11), 2031-2044.

Cisco. (2009). Cisco visual networking index: Forecast and methodology, 2009-2014. 2009-2014. Retrieved April 2, 2011 from http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf.

Clark, J. (2010). *Tapworthy: Designing great iPhone apps*. Beijing: O'Reilly Media.

Dempsey, L. (2009, January 5). Always on: Libraries in a world of permanent connectivity. First Monday, 14. Retrieved April 2, 2011 from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2291/

[8] iPod Touch 4th generation, iPad 2nd generation and devices with camera equipment.

Facebook. (2011). Press room: Statistics. Retrieved April 2, 2011 from http://www.facebook.com/press/info.php?statistics

Fling, B. (2009). *Mobile design and development: Practical concepts and techniques for creating mobile sites and Web apps*. O'Reilly.

Foresman, C. (2009). Apple responsible for 99.4% of mobile app sales in 2009 (Updated). Retrieved April 2, 2011 from http://arstechnica.com/apple/news/2010/01/apple-responsible-for-994-of-mobile-app-sales-in-2009.ars

Group, N.N. (2009). Usability of mobile websites. Retrieved April 2, 2011 from http://www.service2media.com/usability-of-mobile-websites/

Hacettepe. (2010). Kütüphanelerimiz. Retrieved April 2, 2011 from http://www.library.hacettepe.edu.tr/BelgeGoster.aspx?F6E10F8892433CFFAAF6AA849816B2EF134E1A1D2562C276

Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). *The 2010 Horizon report*. Austin, Texas: The New Media Consortium.

Klein, G. (1999). *Sources of power: how people make decisions*. MIT Press.

Krug, S. (2006). *Don't make me think!: a common sense approach to web usability*. New Riders.

Kukulska-Hulme, A. (2007). Mobile usability in educational contexts: What have we learnt? The International review of research in open and distance learning (2). Retrieved April 2, 2011 from http://www.irrodl.org/index.php/irrodl/article/viewArticle/356/879

Morgan Stanley. (2009). *The Mobile Internet report: Ramping faster than desktop Internet, the mobile Internet will be bigger than most think*. Retrieved April 2, 2011 from http://www.morganstanley.com/institutional/techresearch/pdfs/Theme_6_Data_Growth.pdf

Morville, P. & Rosenfeld, L. (2006). *Information architecture for the World Wide Web*. Sebastopol, CA: O'Reilly.

Needham, G. & Ally, M. (2008). *M-libraries: Libraries on the move to provide virtual access*. London: Facet.

Nielsen, J. (1993). *Usability engineering*. Boston, MA: Academic Press.

Nielsen, J. (2009). Mobile usability. Retrieved April 2, 2011 from http://www.useit.com/alertbox/mobile-usability.html

Owoh, B. (2010). A Quick look at mobile Web designs. Retrieved April 2, 2011 from http://sixrevisions.com/user-interface/a-quick-look-at-mobile-web-designs/

Rainie, L. & Anderson, J. Q. (2008). Pew Internet and American life project: The Future of the Internet III. Retrieved May 28, 2011 from http://www.pewinternet.org/Reports/2008/The-Future-of-the-Internet-III.aspx

Rayport, J. F. (2009). Social networks are the new Web portals. Retrieved April 2, 2011 from http://www.businessweek.com/technology/content/jan2009/tc20090121_557202_page_2.htm

Uçak, N. Ö. & Çakmak, T. (2009). Web sayfası kullanılabilirliğinin ölçülmesi: Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü web sayfası örneği. *Türk Kütüphaneciliği,* 23(2), 278-298.

Semantic Search in a Digital Library with Bulgarian Folk Songs

Maria M. Nisheva-Pavlova *and* **Pavel I. Pavlov**

Faculty of Mathematics and Informatics, Sofia University, 5 James Bourchier blvd., Sofia, Bulgaria.
Email: marian@fmi.uni-sofia.bg, pavlovp@fmi.uni-sofia.bg

Abstract: *The paper presents some aspects of an ongoing project aimed at the development of technologies for digitization of Bulgarian folk music and building a heterogeneous digital library with Bulgarian folk songs presented with their music, notes and text. This digital library will provide both digital preservation of the sound recordings, lyrics and notations of Bulgarian folk songs and a possibility for new interpretations of the archaic Bulgarian folklore heritage. Some facilities of the search engine under development to implement various types of search and access to the library resources are analyzed in the paper. The emphasis of the discussion falls on the tool provided for semantic search in the lyrics of songs.*

Keywords: *Digital library; metadata; semantic web; ontology; search engines.*

Introduction

Bulgarian folk music is a valuable resource of cultural memory and is among the main characteristics of the national identity of Bulgarian people. Throughout the years the Bulgarian researchers of musical folklore have written down hundreds of thousands of musical folk samples (songs and instrumental melodies) in lyrics and notes. Part of these music notations has been published, another part is preserved as manuscripts in specialized institutional or personal archives.

The paper presents some initial results of the activities within an ongoing project aimed at the development of technologies for digitization of Bulgarian folk music and building a digital library (named DjDL) with Bulgarian folk songs presented with their text, notes and music. DjDL will serve as a platform for digital preservation of the sound recordings, lyrics and notations of Bulgarian folk songs which will provide a possibility for exploration and new interpretations of the archaic Bulgarian folklore heritage. The main facilities of the search engine under development to realize various types of search and access to the lyrics of songs are discussed in detail.

Main Characteristics of the Library Resources

Currently DjDL keeps a collection of digital objects which represent a part of the unpublished archive manuscripts of Prof. Todor Dzhidzhev containing recordings of over 1000 folk songs from the Thracia region of Bulgaria. Completion and diversification of this collection with music from various folk musical dialects has been planned for the near future.

After a careful examination, LilyPond[1], extended with tools for using some specific symbols and other elements necessary for notation of the Bulgarian national music (Kirov, 2010), was chosen as a technological base for coding of note records from original manuscripts with field recordings of Bulgarian folk songs.

[1] LilyPond ... music notation for everyone, http://lilypond.org/

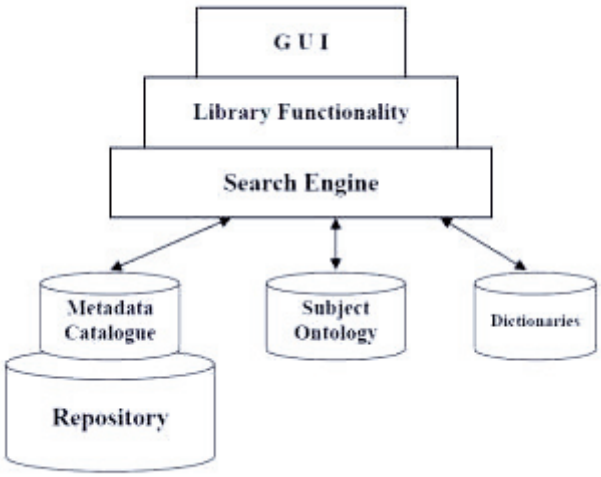

Figure 1. Functional structure of DjDL

DjDL has the typical architecture of an academic digital library with heterogeneous resources. Its functional structure is shown in Figure 1.

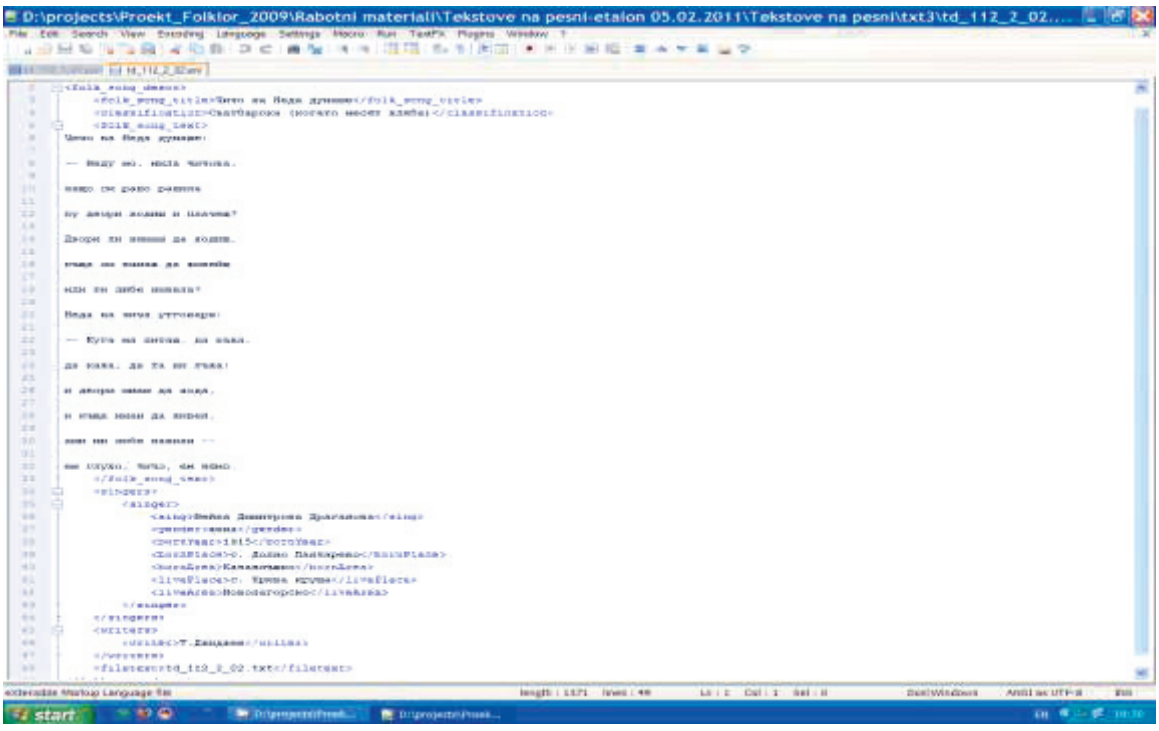

Figure 2. Part of a catalogue description

The library catalogue contains short descriptions (in XML format) of the songs included in the repository. Various types of relevant metadata (metadata attributes) are provided, for example: the title of the song, the song genre according to different classification schemes (e.g. according to the typical time and space of performance, thematic focus(es), context of performance, etc.), the region of folk dialect, data specifying the informant (the person who conveyed the song to folklorists) and the time and place of gathering the song (the folklorist who gathered the song, the singer(s), the time and

place of record, etc.). More precisely, each catalogue entry contains the text (i.e., the lyrics) of a particular song accompanied by the corresponding metadata.

Figure 2 displays a selected part of the catalogue description of the folk song entitled "Чичо на Неда думаше" ("Neda's uncle speaking to her").

The repository of DjDL contains heterogeneous resources of the following types (Peycheva et al., 2010):

- lyrics of songs (in PDF format);
- notations of songs (in LilyPond and in PDF format);
- musical (MP3) files with the authentic performances (as far as such exist in the archives) of the songs;
- musical (MIDI) files generated with the use of LilyPond from the notations of the songs.

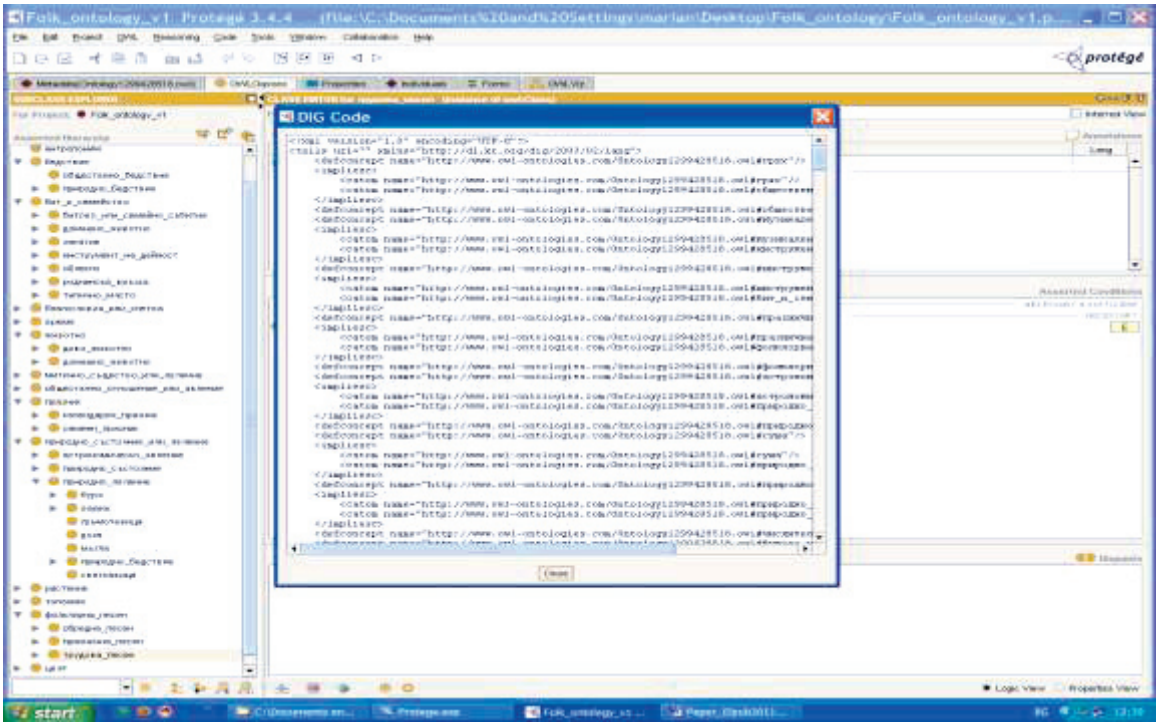

Figure 3. Part of the concept hierarchy and the corresponding DIG code

The subject ontology consists of several interrelated subontologies needed by the search engine of DjDL and developed especially for the purposes of the discussed project:

- ontology of folk songs – includes various genre classifications of folk songs (by their thematic focus – historical, mythical, etc.; by the context of performance – Christmas folk songs, harvest songs, etc.; by their cultural functions – blessing, oath, wooing, etc.);
- ontology of family and manner of life;
- ontology of impressive events and natural phenomena;
- ontology of social phenomena and relationships;
- ontology of mythical creatures and demons;
- ontology of settlements.

For the development of the subject ontology we used one of the most popular ontology editors – Protégé/OWL (Knublauch, 2003; Knublauch et al., 2004). The implementation of the search engine is especially oriented to the corresponding DIG code with two simple automatically made changes in it: reduction of the ontology URL in all class, property and individual names and replacement of the underscore characters with spaces.

Figure 3 shows a part of the "is-a" hierarchy of concepts included in the particular subontologies.

The search engine of DjDL also uses some proper dictionaries available in digital (XML) format – a dictionary of synonyms and a dictionary of obsolete and dialect words.

Functionalities of the Search Engine

The search engine of DjDL supports two main types of search: keywords-based and semantic search. Its current version realizes some facilities for search in the catalogue metadata and the lyrics of songs only. The design and the implementation of this search engine are based on some former results of the authors (Nisheva-Pavlova & Pavlov, 2010) and some ideas from de Juan & Iglesias (2009) and McGuinness (2003). Its full palette of functionalities has been specified after a careful study of the requirements of the typical user groups (specialists and researchers in ethnomusicology, verbal folklore and folkloristics in general, philologists, etc.).

The user queries define restrictions on the values of certain metadata attributes and/or the lyrics of the required folk songs. The search procedure implies the performance of some pattern matching activities in which the catalogue descriptions are examined one by one and those having a specific set of element values that match the corresponding components of the user query are marked in order to form the search result.

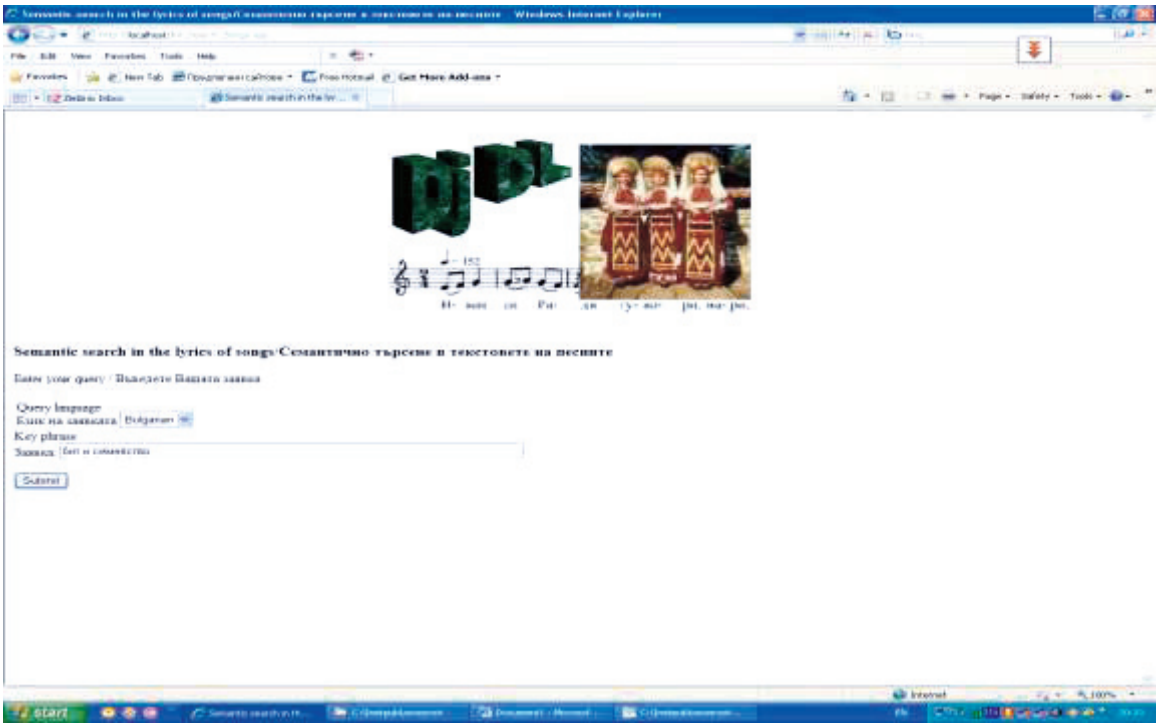

Figure 4. A sample user query for semantic search

The matching process within the keywords-based search consists in testing the appropriate sources for equality.

As a first step in the construction of a query for keywords-based search, the user is asked to indicate the search source(s) – search in the lyrics of songs, search in the metadata (in the catalogue descriptions) or combined search in the lyrics of songs and catalogue metadata. One may define a search query consisting of an arbitrary number of words or phrases as well as specify proper logical connectives between them: conjunction (and) or disjunction (or). Negation (not) is also allowed as a unary operator indicating that the negated word or phrase should not appear in the corresponding text. As a result of the user query processing, a list of links to the discovered files with lyrics of songs is properly displayed. This list may be ordered by the titles of songs or by the number of appearances of the words (or phrases) included in the user query.

The following list contains some typical examples of queries for keywords-based search:

- search (and retrieval) of songs whose lyrics contain specific words or phrases;
- search of songs with distinct thematic focus or context of performance;
- search of songs performed by a given singer;
- search of songs performed by singers from a given place.

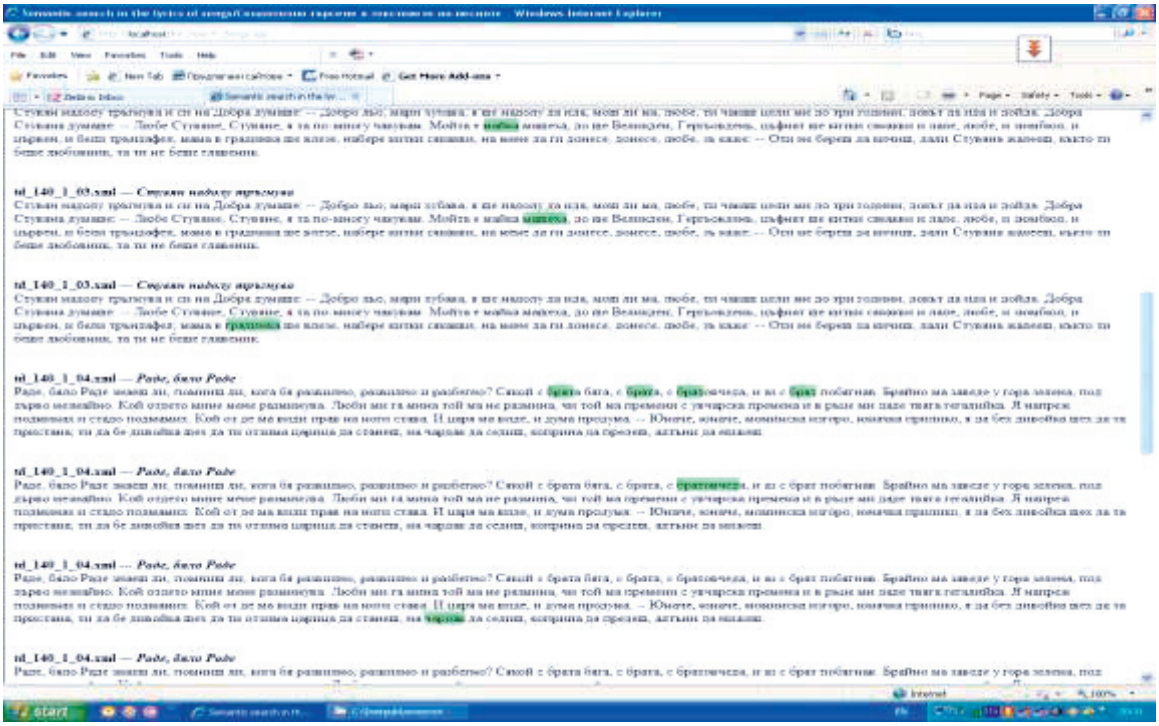

Figure 5. Part of the search results for a user query containing the phrase “popular customs and family”

The semantic search service in DjDL is aimed at the provision of some additional functional facilities for extension, generalization and refinement (automatic reformulation according to the available explicit domain knowledge) of the queries for keywords-based search.

The extension (more precisely, the augmentation) of the user query is based on the use of the subject ontology and the available dictionaries – the dictionary of synonyms and the dictionary of obsolete and dialect words. First of all, an exhaustive breadth-first search in the graph representing the “is-a” concept hierarchy described by the subject ontology is performed, starting from the node which corresponds to the original user query. The names of the visited nodes, i.e. the corresponding more specific concepts from the ontology (the concept hierarchy) are added to the one given by the user. The names of the existing instances of these concepts are added as well. Then the search engine adds to the newly constructed set of queries some synonyms, derivatives, dialect and obsolete forms of the main terms found in the mentioned dictionaries. Thus the user query is augmented as far as possible and in fact has the shape of a disjunction of all included forms of concepts and instance names. In this form it is ready for further processing.

Figure 4 illustrates a sample user query for semantic search containing the phrase “popular customs and family” (“бит и семейство” in Bulgarian).

Figure 5 shows a screenshot displaying part of the search results for this query which contains the titles of some retrieved songs (the text in bold italic type) and their texts in a compressed form as well as the corresponding file names (the names of the corresponding catalogue entries – the text in bold face). The discovered words and phrases that are semantically related to the original user query are highlighted.

A careful examination of the search results displayed in Figure 5 shows that some of the highlighted words denote concepts which are more particular than the query one (in accordance with the "is-a" hierarchy described within the subject ontology) – typical scenes of narrative in folk songs („чардак") and ties of relationship („майка", „мащеха", „брат", „братовчед"). The word „градинка" is a diminutive form of „градина" ("garden") which is also a typical scene of narrative.

As examples of queries for semantic search of interest for folklorists (according to Peycheva & Grigorov, 2010) that can be executed by the search engine of DjDL, one may indicate the queries for search and retrieval of:

- songs devoted to important historical events or social phenomena;
- songs in which exciting natural or astronomical phenomena are described or mentioned;
- songs in which typical (or typical for a certain region) folk beliefs are described;
- songs in which elements of country work and life are described or mentioned;
- songs in which important family events or typical family relationships (daughter-in-law – mother-in-law, son-in-law, mother-in-law, etc.) or joyful/unfortunate family events are mentioned.

The current implementation of the search engine also provides some facilities for processing of user queries containing examination of equality or inequality. For example, it is possible to formulate and execute queries for search of:

- songs performed alone/in a group;
- songs performed by men/women only;
- songs performed by a particular singer (grouped by the names of singers);
- songs performed by singers born in a particular settlement or region (grouped by settlements/regions);
- songs performed by singers who have moved from a particular settlement or region;
- songs performed to the west/east/north/south of a specific settlement/region;
- songs performed in a specific region (grouped by regions of performance);
- songs in which at least/more than/exactly a specific number of toponyms/anthroponyms are mentioned.

In contrast with our former experience in implementation of similar techniques for the purposes of semantic search in collections of digitized mediaeval manuscripts (Pavlov & Nisheva-Pavlova, 2006) and archival documents (Nisheva-Pavlova et al., 2007), the results obtained in the discussed case of search in the texts of folk songs may be evaluated as insufficient. We suppose that the reason for this conclusion lies in the vast use of similes, metaphors, idioms and other sophisticated or language-dependent stylistic devices in the folklore lyrics. With the aim of overcoming some aspects of this problem, we intend to design and use a set of proper patterns of typical stylistic or thematic constructs which could be matched with relatively large parts of the texts of folklore songs. For example, we have already defined a number of patterns of constructs standing for "unfaithfulness", "jealousy", "discontent" and "sedition" as well as the corresponding pattern matching rules and recently performed various experiments with them.

Our current activities are directed to the design and implementation of a tool for flexible and convenient (intuitive) construction of complex user queries (without using natural language), in particular queries including proper combinations of the example types mentioned above.

The next step will be to extend the functional facilities of the search engine of DjDL with a proper tool for semantic search and knowledge discovery in the notes of songs. A main goal in this direction will be to automate the further study of some musical characteristics of Bulgarian folk songs (e.g., their melodies and rhythms) with the aim of discovering similarities of songs according to various criteria.

Summary and Future Work

The final version of DjDL will be developed with the aim of providing a complete set of tools which will be useful for a series of further studies in folkloristics, philology and musicology, for example:

- creation of a set of the phonetic variations which every sound can undergo in accordance with the Bulgarian lexicology and dialectology;
- classification of folk songs according to different criteria: technical, implicit-musical, implicit-textual, statistical, genre, cultural, etc.;

- creation of frequency dictionary and concordance, based on the lyrics of songs and inferring of theoretical conclusions from the obtained results.

Acknowledgements

The project "*Information technologies for presentation of Bulgarian folk songs with music, notes and text in a digital library*" has been supported by the Bulgarian National Science Fund under Grant DTK 02/54.

References

De Juan, P. & Iglesias, C. (2009). Improving searchability of a music digital library with semantic web technologies. *Proceedings of the 21st International Conference on Software Engineering & Knowledge Engineering (SEKE'2009)* (pp. 246-251). Boston, Massachusetts: Knowledge Systems Institute Graduate School, ISBN 1-891706-24-1.

Kirov, N. (2010). Digitization of Bulgarian folk songs with music, notes and text. To appear in: *Review of the National Center for Digitization*, ISSN 1820-0109. Retrieved, April 3, 2011, from http://eprints.nbu.bg/627/1/NKirov_folk.pdf.

Knublauch, H. (2003). An AI tool for the real world: Knowledge modeling with Protégé. *JavaWorld*, June 20.

Knublauch, H., Fergerson, R., Noy, N., & Musen, M. (2004). The Protégé OWL plugin: An open development environment for semantic web applications. In S.A. McIlraith et al. (Eds.). The Semantic Web: ISWC 2004. *Lecture Notes in Computer Science*, 2004, Volume 3298/2004, 229-243. Heidelberg: Springer.

McGuinness, D. (2003). Ontologies come of age. In D. Fensel, J. Hendler, H. Lieberman, & W. Wahlster (Eds.), *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. Cambridge, MA: MIT Press, ISBN 978-0-262-06232-9.

Nisheva-Pavlova, M., Pavlov, P., Markov, N., & Nedeva, M. (2007). Digitisation and access to archival collections: A case study of the Sofia Municipal Government (1878 – 1879). L. Chan and B. Marten (Eds.) *ELPUB2007. Openness in Digital Publishing: Awareness, Discovery and Access - Proceedings of the 11th International Conference on Electronic Publishing held in Vienna, Austria 13-15 June 2007* (pp. 277-284). Vienna: ÖKK-Editions, ISBN 978-3-85437-292-9.

Nisheva-Pavlova, M. & Pavlov, P. (2010). Search engine in a class of academic digital libraries. In T. Hedlund & Y. Tonta (Eds.), *Publishing in the Networked World: Transforming the Nature of Communication. 14th International Conference on Electronic Publishing 16-18 June 2010, Helsinki, Finland* (pp. 45-56). Helsinki, Edita Prima Ltd, ISBN 978-952-232-085-8.

Pavlov, P. & Nisheva-Pavlova, M. (2006). Knowledge-based search in collections of digitized manuscripts: First results. In Bob Martens and Milena Dobreva (eds.). *Digital Spectrum: Integrating Technology and Culture - Proceedings of the 10th International Conference on Electronic Publishing. Bansko, Bulgaria June 14-16, 2006* (pp. 27-35), Sofia: FOI-Commerce, ISBN 978-954-16-0040-5.

Peycheva, L. & Grigorov, G. (2010). How to digitalize folklore song archives? To appear in: *Review of the National Center for Digitization*, ISSN 1820-0109.

Peycheva, L., Kirov, N., & Nisheva-Pavlova, M. (2010). Information technologies for presentation of Bulgarian folk songs with music, notes and text in a digital library. *Proceedings of the Fourth International Conference on Information Systems and Grid Technologies (ISGT'2010)* (pp. 218-224). Sofia: St. Kliment Ohridski University Press, ISBN 978-954-07-3168-1.

Mate in 3D – Publishing Interactive Content in PDF3D

Frank Breuel
Universität Rostock, Germany and Fraunhofer Austria. Email: frank.breuel@uni-rostock.de

René Berndt
Graz University of Technology, Austria. Email: r.berndt@cgv.tugraz.at

Torsten Ullrich
Fraunhofer Austria Research GmbH, Graz, Austria. Email: torsten.ullrich@fraunhofer.at

Eva Eggeling
Fraunhofer Austria Research GmbH, Graz, Austria. Email: eva.eggeling@fraunhofer.at

Dieter W. Fellner
Fraunhofer IGD and TU Darmstadt, Germany. Email: d.fellner@igd.fraunhofer.de

Abstract: *In this paper we describe a pipeline for publishing interactive multimedia content. The Portable Document Format (PDF) offers the possibility to include 3D visualizations, textual representation and interactivity (via scripting technology) in one multimedia container, which will be the upcoming standard for multimedia long-term archiving. Our system demonstrates its potential for surplus value eBooks. By the example of chess we developed a publishing pipeline to create interactive books. Usually, chess games and positions are recorded using the algebraic chess notation, which is mainly an annotated list of moves. In combination with a time-dependent 3D visualization, each move corresponds to a specific game position. This correspondence is encoded in hyperlinks from the textual representation to the 3D visualization. This linkage improves readability and usability of chess notations significantly. Furthermore, using an established file format our eBooks can be opened by any compliant PDF viewer.*

Keywords: *PDF3D; chess; multimedia; interactive content.*

Introduction

According to José Raúl Capablanca[1] -a great world chess master- "Chess books should be used as we use glasses: to assist the sight, although some players make use of them as if they conferred sight" (Evans & Wiswell, 1953). Chess, the popular board game, has been fascinating its players for centuries. Due to its tactical character, the analysis of ancient games of chess is an important aspect in chess theory. Many games are recorded, evaluated and analyzed by interested players. Usually, chess games and positions are recorded using the algebraic chess notation (FIDE, 2008), which is mainly an annotated list of moves. In order to analyze a game, one has to take a chess board and chess players and replay the game move by move; especially the analysis of variants and alternatives can be cumbersome.

The success story of information technology in the field of chess applications comprehends an algorithmic breakthrough best described by two quotes of Garry Kasparov[2], a famous world chess champion: "Inevitably the machines must win, but there is still a long way to go before a human on his or her best day is unable to defeat the best computer."[3] Only a few years later he said that "ironically, the main task of chess software companies today is to find ways to make the program weaker, not stronger, and to provide enough options so that any user can pick from different levels

[1] http://en.wikipedia.org/wiki/Jos%C3%A9_Ra%C3%BAl_Capablanca
[2] http://en.wikipedia.org/wiki/Kasparov
[3] http://www.chessquotes.com/player-kasparov

and the machine will try to make enough mistakes to give him a chance."[3] Despite this algorithmic breakthrough, the questions of usability and long-term archiving have been unanswered so far.

In this paper we present a solution for digital publication and archiving of chess games based on established PDF standards. The resulting documents can be read with any PDF viewer. Our system is capable of generating an interactive, self-contained PDF document automatically from a given portable game notation (PGN) file (Edwards, 1994). The PGN file format is the most popular way to store chess games digitally. Thousands of played games in different varieties have been made available up to now in this format.

PDF is the perfect choice for publishing multimedia eBooks. The next version PDF/A-2 (ISO 2008) of the archival standard PDF/A (ISO, 2005) based on PDF 1.7 is expected to contain support for 3D. Consequently it becomes possible to generate state-of-the-art multimedia content that can be used for long-time archiving and storage. Furthermore, chess games that were created with our software can thus be displayed at any time by any PDF viewer. Using digital library techniques, every game of chess can be identified uniquely. Embedded into a standardized document, a chess game can be identified via the document's digital object identifier (DOI) and thus be referenced.

Related Work

Concerning chess books several techniques for publication are established. The most commonly used chess books (tournament publications as well as textbooks) are published the classical way (Appelt, 1988).

With upcoming information technology systems, a growing number of chess publications have been digitally born. These systems can be distinguished from each other by the visualization technique used. On the one hand, there are a large number of software packages that need to be installed locally. On the other hand, online solutions on various websites allow a user to recapitulate a chess game using advanced web technology, namely JavaScript and HTML.

While there are a variety of eBooks formats[4] available, only a few offer support for multimedia content. This is not surprising since the "classic" eBooks formats mostly target hardware devices, which are not designed for power-consuming multimedia.

Software products that have to be installed locally, such as *Chess Base Light,* offer a range of options for visualizing and editing chess games. Such programs usually supply a big chess database including various matches. In addition, they offer the possibility to recapitulate a chess match graphically with included annotations, comments, etc. Furthermore, it is possible to create and save new games. Nevertheless, these applications are highly dependent on the operating system, cannot be used on eBook devices and are not a serious option for long-term archiving.

The online approaches via current web technology show an uneven picture. Some solutions need permanent online access, some have limited visualization capabilities, some are "optimized" for special browsers and environments, and none of them is a serious alternative for long-term archiving. The new features of HTML5 (W3C, 2011), like the "canvas" element for 2D drawings, promise a new and cross-platform solution for creating and presenting interactive content. Even 3D support will be available through WebGL (Khronos Group, 2011). Currently only Mozilla Firefox and Chrome (and the beta releases of Safari and Opera) offer support for WebGL. HTML5 will become a W3C recommendation in 2014.

Technology

PDF 3D

With the introduction of PDF 1.6 (Adobe Acrobat 7) it is possible to add 3D content to a PDF document. The 3D content – also referred to as 3D annotations – is defined within a rectangular area of the page, which optionally can be overlaid with a bitmap for non-3D capable PDF viewers. Basic features of 3D annotations include multiple pre-defined views, various render modes and different navigation modes.

[4] http://en.wikipedia.org/wiki/Comparison_of_e-book_formats

A crucial aspect of the interactive 3D annotation is the definition of the behavior for the scene. For this purpose Adobe introduced the 3D JavaScript API (Adobe, 2007). This API is separate from the document-level JavaScript API for manipulating the "classical" PDF content. This API provides special methods and classes to realize 3D programming, e.g., *Node*, *Light*, *Material* or *Camera*. Callback classes for various events can be used to define a custom interaction for the 3D scene, e.g. mouse events, keyboard events, timer, etc.

According to the standard by Adobe there is support by two formats for representing 3D content in a PDF document. These formats are the RPC file format (Adobe, 2008) and the Universal 3D File Format (U3D) (ECMA International, 2007).

XSL-FO (XSL-Formatting Objects)

XSL Formatting Objects (XSL-FO) is an XML-based markup language for formatting and presentation of XML-documents (W3C, 2006). It defines a set of elements and attributes for describing the visual appearance of XML-encoded data.

The general idea behind XSL-FO is to transform an XML document containing data into another XML document containing the data and the XSL-FO elements for the visual presentation. This document is then converted by an FO processor (FOP) to a given target format, which has to be printable, writeable or even both. XSL-FO is designed to support various render targets depending on the implementation of the FO processor. Nevertheless, in most cases the render target is PDF.

FO3D (Formatting Objects 3D)

FO3D is an extension to the XSL Formatting Objects (XSL-FO) standard (Buchgraber, Berndt, Havemann, & Fellner, 2010). Since the XSL-FO is designed for rendering two dimensional layouts, the FO vocabulary lacks support for embedding 3D content. However, support for arbitrary objects is already included within the XSL standard.

Figure 1 shows the definition of a 3D object using FO3D. The root element of the FO3D object with 3D content is *object-3d*. The following sub-tree contains information about additional data used for creating the 3D annotation, like additional resources (e.g., other 3D models, images) or JavaScript code.

FO3D utilizes the *fo:instream-foreign-object* element, which "is used for an inline graphic or other 'generic' object where the object data resides as descendants of the *fo:instream-foreign-object*, typically as an XML element sub tree in a non-XSL namespace" (W3C 2006). FO3D primarily targets the integration of 3D content and related metadata in PDF documents; therefore the proposed FO3D vocabulary contains various settings and concepts following the ISO-standardized PDF-1.7 specification (ISO-32000-1 2008). A detailed description of the FO3D vocabulary can be found in Buchgraber et al. (2010). Figure 2 shows the typical workflow for XSL-FO processing with FO3D. FO3D has been implemented for Apache FOP (Apache, 2011) and is available for download at http://fo3d.sourceforge.net.

```
1   <?XML version="1.0" encoding="UTF-8"?>
2   <fo:root XMLns:fo=http://www.w3.org/1999/XSL/Format
    XMLns:fox="http://XMLgraphics.apache.org/fop/extensions">
3   <fo:layout-master-set>
4   …
5   <fo:instream-foreign-object XMLns="http://fo3d.sourceforge.net/ns">
6      <object-3d src="./Fop files/Models/block.u3d" width="15cm" height="15cm"
        altsrc="./Fop files/Models/Overlay.jpg">
7         <default-view background-color="white" render-mode="solid" light-mode="cad">
8            <camera position="0,50,0" target="0,0,0" up="0,0,-1" />
9         </default-view>
10        <resources>
11           <resource src="./Fop files/Models/Chessfigures/u3d/pawnBlack.u3d"
              name="pawnBlack" />
12        </resources>
13        <scripts>
14           <script type="open" src="./Fop files/Sources/Board.js" />
15           …
16        </scripts>
17     </object-3d>
18  </fo:instream-foreign-object>
19  …
```

Figure 1. Embedded 3D content information in an FO file. The object-3d builds the root of a 3D annotation.

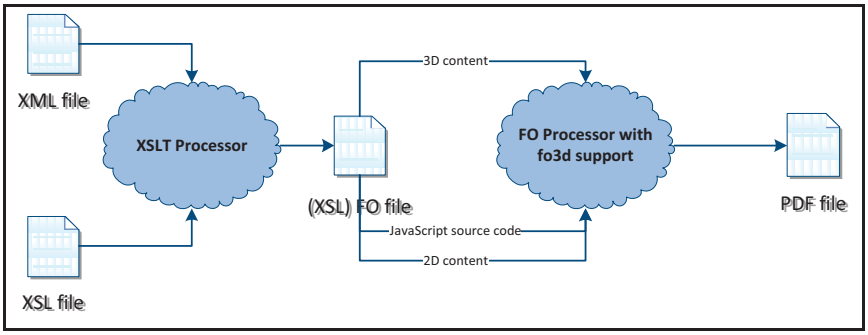

Figure 2. Workflow diagram of the XSL-FO formatting process, embedding 3D content in PDF documents (Buchgraber, et al., 2010)

Processing Pipeline

Our proposed system implements the complete processing pipeline starting from the source PGN to the final PDF document (see Figure 3). Our system consists of several parts that will be explained in more detail in the following paragraphs.

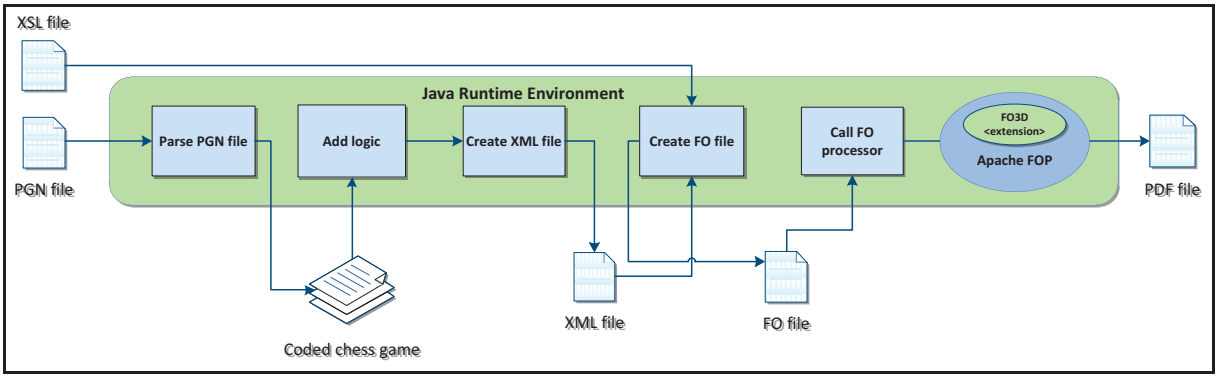

Figure 3. General structure of the processing pipeline

Parsing a PGN file

The first step of our pipeline is to parse the given PGN file. The PGN standard was developed and published by Steven J. Edwards in 1994 (Edwards, 1994). It stores a game notation as plain text, which can be very easily read and written by humans (Figure 4). Also, most chess programs, e.g. ChessBase[5], can export their data to PGN. The first part of a PGN file consists of a set of tuples containing general metadata of the match such as date, player information, location, etc. enclosed in square brackets.

```
[Event "New York Rosenwald"]
[Site "New York"]
[Date "1956.10.07"]
[White "Byrne,Donald"]
[Black "Fischer,Robert James"]
[Result "0-1"]
[Eco "D97"]

1.Nf3 {A noncommittal move. From here, the game can develop into a number of different
openings.} Nf6 2.c4 g6 3.Nc3 Bg7 4.d4 O-O {Fischer castles, concentrating on protecting his
king immediately.} 5.Bf4 d5 6.Qb3 {The so-called Russian System, putting pressure on
Fischer's central d5 pawn.} dxc4 7.Qxc4 c6 8.e4 Nbd7  9.Rd1 Nb6 10.Qc5 Bg4 {At this point,
Byrne's pieces are more developed, and he controls the center squares. However, Fischer's
king is well-protected, while Byrne's king is not.} 11.Bg5 (11. Be2 Nfd7 12. Qa3 Bxf3 13.
Bxf3 e5 14. dxe5 Qe8 15. Be2 Nxe5 16. O-O {and white is better.}) Na4 12.Qa3 {Here Fischer
cleverly offers up his Knight, but if Byrne takes it with} (12.Nxa4 {Fischer will play}
Nxe4 {, and Byrne then suddenly has some terrible choices:} 13. Qxe7 (13. Bxe7 Nxc5 14.
Bxd8 Nxa4 15. Bg5 Bxf3 16. gxf3 Nxb2 {gives Fischer an extra pawn and ruins Byrne's pawn
structure.})
…
```

Figure 4. PGN file including moves, comments and variants in the chess match

Next up follows the declaration of all moves made in that game. In addition, a move can contain comments (enclosed in curly brackets), NAGs (Numeric Annotation Glyphs) and alternative variants (enclosed in round brackets). NAGs (Edwards, 1994) are used to annotate chess games providing an assessment of a chess move or a chess position, e.g. good move, poor move, white has a moderate advantage, etc.

Our PGN parser builds up a tree structure based on half-moves. A half-move is defined as the move of one player (the half-move of white followed by the half-move of black is then a full move). Each node in the tree represents a half-move H_1. The next half-move H_2 is stored as the first child of the node containing H_1. Variants are stored as siblings to H_2. A match without variants would result in a linear graph. Each node contains several pieces of information about the half-move, like the color, the number of already made half-moves and the source and destination field of the piece.

Adding additional data

In most cases the abbreviated (or short) algebraic notation is used. It just records the moving piece along with the destination field. If this information does not unambiguously define the half-move, information about the source field is added, e.g. Nf3d or Qxe5. Here the first notation means that the *Knight* on field *f3* has to move to field *d*; the second signifies that the *Queen* beats the chess player located on field *e5*. There exist a lot of possibilities to quote a single chess move and all of them have to be parsed and checked.

Therefore the starting-field and the actually moved chess piece will be calculated for each half-move. If there are more than one or even no possible chess piece for that half-move notation, the semantic structure of the file seems to be incorrect so the file will be rejected. Otherwise, if exactly one chess piece was found, this information will be added to the node containing the half-move, and the position of all pieces will be updated. These calculations have to be performed to the main line and all variants.

[5] http://www.chessbase.com/

Generating the XML and (XSL) FO file

The complete parsed half-move tree is then exported as an XML document. This XML document is used for generating the 3D JavaScript code for that particular match.

Figure 5 shows the structure of the XML document. The first part contains the metadata of the match, as found in the PGN file. The match itself is listed as a sequence of half-moves. Every outlined half-move contains a unique identification number, a reference to the previous and to the next move.

```
1   <?xml version="1.0" encoding="UTF-8"?>
2   <chessgames>
3   <chessgame>
4   <filename>gameOfTheCentury.pgn</filename>
5   <result>0-1</result>
6   <attributes>
7      <attribute>[Event "New York Rosenwald"]</attribute>
8      …
9   </attributes>
10  <moves>
11     <move>
12        <halfmoveWhite moveNumber="1" mainLine="true" color="white" moveFull="Nf3"
           id="0" idParent="-1" idChild="1">
13          <chessmove from="g1" to="f3" switch=" " />
14          <comment>A noncommittal move. From here, the game can develop into a
             number of different openings.</comment>
15        </halfmoveWhite>
16        <halfmoveBlack moveNumber="1" mainLine="true" color="black" moveFull="Nf6"
           id="1" idParent="0" idChild="2">
17          <chessmove from="g8" to="f6" switch=" " />
18        </halfmoveBlack>
19     </move>
20  <move>
21  …
```

Figure 5. XML document example containing the sequence of half-moves.

Currently there is no support for adding a 2D overlay to the 3D annotation, e.g. to display the comments for a particular half-move. Our solution is to use billboarding, i.e. rendering the text to a bitmap and including this bitmap as a resource within the 3D annotation. Figure 6 shows how the images for the comments are represented within the XML document.

```
1   <?xml version="1.0" encoding="UTF-8"?>
2   <chessgames>
3      <chessgame>
4         …
5         <moves>
6         …
7         </moves>
8         <commentImages>
9            <commentImg commentId="commentImage0">./Fop
              files/Models/Images/commentImage_0.jpg</commentImg>
10           <commentImg commentId="commentImage5">./Fop
              files/Models/Images/commentImage_5.jpg</commentImg>
11           …
12        </commentImages>
13        <moveImages>
14           <moveImg moveId="moveImage0">./Fop files/Models/Images/moveImage_0.jpg
             </moveImg>
15           …
16        </moveImages>
17     </chessgame>
18  </chessgames>
```

Figure 6. Associated relative paths for each comment and move in the XML file.

The resulting XML document is then transformed using XSL into an FO document. Figure 7 shows how a half-move is laid out for the PDF document. Line 25 shows how the textural representation of the half-move can interact with the 3D annotation.

```
<?XML version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0" XMLns:xsl="http://www.w3.org/1999/XSL/Transform">
…
<xsl:template match="move">
   <inline XMLns="http://www.w3.org/1999/XSL/Format">
      <xsl:if test="halfmoveWhite/@mainLine='true'">
         <inline text-decoration="underline">
            <xsl:if test="halfmoveWhite/@moveNumber != -1">
               <xsl:value-of select="halfmoveWhite/@moveNumber" />
               <xsl:text>. </xsl:text>
            </xsl:if>
         </inline>
      </xsl:if>
      <xsl:if test="halfmoveWhite/@mainLine='false'">
         <xsl:if test="halfmoveWhite/@moveNumber != -1">
            <xsl:value-of select="halfmoveWhite/@moveNumber" />
            <xsl:text>. </xsl:text>
         </xsl:if>
      </xsl:if>
   </inline>

   <xsl:variable name="numberWhite">
      <xsl:value-of select="halfmoveWhite/@id" />
   </xsl:variable>
   <xsl:variable name="javascriptWhite">
      <xsl:value-of select="concat('javascript:var
        h=getAnnots3D(0)[0];if(h.activated)h.context3D.goTo(',$numberWhite,')')" />
   </xsl:variable>
…
```

Figure 7. XSL document example, containing information on how to layout the chess data

Creation of the PDF document

The generated FO document is then processed by Apache FOP. In order to process the custom chess XML structure described above, FO3D has been extended using a custom FO3D extension. FO3D custom extensions are typically used for tasks which are not possible using XSLT and Javascript, e.g. generating images or loading arbitrary external files referenced from the XML input.

Figure 8 shows the use of the PGN extension within the FO document. The PGN extension generated the necessary JavaScript code, which is necessary for the animation of the chess match within the 3D annotation. It also provides a mechanism for the document level JavaScript to interact with the 3D annotation.

```
<?xml version="1.0" encoding="UTF-8"?>
<fo:root xmlns:fo=http://www.w3.org/1999/XSL/Format
xmlns:fox="http://xmlgraphics.apache.org/fop/extensions">
   <fo:layout-master-set>
      <fo:simple-page-master master-name="A4"
         …
            <extensions>
               <pgn src="./pgnAsXml.xml" />
            </extensions>
         </object-3d>
      </fo:instream-foreign-object>
   </fo:block>
   <fo:block font-family="Helvetica" font-size="11">
      <inline xmlns="http://www.w3.org/1999/XSL/Format">
         <inline text-decoration="underline">1. </inline>
      </inline>
      <inline font-style="italic" xmlns="http://www.w3.org/1999/XSL/Format">
         <basic-link external-destination="javascript:var
          h=getAnnots3D(0)[0];if(h.activated)h.context3D.goTo(0)" color="blue">
            <inline text-decoration="underline">Nf3</inline>
         </basic-link>
         …
```

Figure 8. Part of the FO file used by the FO processor with the PGN extension.

It is possible to jump to any chess move, to take a step forward or backwards as well as to display comments and the move in the scene itself. Besides the 3D annotation the game notation itself is also added to the PDF document in textual hyperlinks. A left click on such a hyperlink, which represents a

single chess half-move, triggers the 3D annotation to display the position after that particular half-move. Lines 16-18 in Figure 8 show how to create such a clickable hyperlink within the text part of the PDF document.

Figure 9. Selected views of the resulting PDF file to a given chess match. **Top left:** 3D annotation showing the position after Fischer's 3rd move. The comment is directly show within the 3D scene. **Top right:** The final position of the match. **Bottom:** The 3D annotation is displayed in a floating window beside the document allowing the viewer to scroll through the text while keeping the 3D annotation visible even on page breaks

Figure 9 shows a resulting PDF document. The user can interact with the 3D scene in several ways. Thus the scene itself can be displayed in a full-screen mode or as a floating window so the user can scroll through the document and always have a view of the match.

Comments and moves are displayed in the 3D annotation as well as in the text. Buttons allow easy navigation within the chess match to go to the next, previous, last, or to the first move in the match or even to start an automatic play-back of the match.

Furthermore it is possible, by clicking on the lower left button, to enable or to disable the visibility of all buttons and the text fields. This is needed in case the scene is located in a floating window and so small that otherwise the match actually is not comprehensible in a graphical way.

Conclusion and Further Work

In this article we have presented the steps and technologies to develop a complete 3D chess eBook using the PDF standard. The pipeline takes care of parsing PGN files (the most commonly-used chess game description file format), extracting content and creating files for the interfaces.

Although several software solutions are already available today, none of them offers or combines the benefits of modern eBook technology (concerning usability and mobility) with the requirements of a digital library (long-term archiving, indexing, referencing). Our solution is platform independent, as only a PDF reader is needed. Furthermore, PDF with 3D support is expected to become the next electronic document file format for long-term preservation (PDF/A-2). As a consequence, our solution is the perfect basis for archiving chess games. In combination with digital object identifier (DOI) techniques a unique identification of each match is possible.

Concerning future work, two important aspects have to be considered. The current solution is optimized for chess games, and meets the special requirements of this field of application. In the future a more general approach will regard a chess game as a specialization of a process. Process documentation including its visualization, metadata (information about the data processed), and paradata (information about the process itself) is an upcoming challenge in a vast range of information technology applications: from civil engineering to medical care via electronic government.

While the first important aspect generalizes the field of application, the second one specializes the resulting documentation format. Different fields of application require adaptations to systems already in use. As our pipeline has a high degree of flexibility and easy expandability, our solution is prepared for cross-media publishing. Consequently, integration tasks should not pose any problems.

Acknowledgments

The work presented in this paper was partially supported by the German Research Foundation DFG under grant INST 9055/1-1 (PROBADO project http://www.probado.de/).

References

Adobe. (2007). *JavaScript™ for Acrobat® 3D Annotations API Reference*. Retrieved April 4, 2011 from http://www.adobe.com/content/dam/Adobe/en/devnet/acrobat/pdfs/js_3d_api_reference.pdf.

Adobe. (2008). *PRC file format specification - Version 8137*. Retrieved April 4, 2011 from http://livedocs.adobe.com.

Apache Software Foundation. (2011). *Formatting Objects Processor – Apache FOP*. Retrieved April 4, 2011 from http://XMLgraphics.apache.org/fop/.

Appelt, W. (1988). Typesetting chess. *The Communications of the TeX Users Group 9.* (pp 284-287).

Buchgraber, G., Berndt, R., Havemann, S., Fellner, D.W. (2010). FO3D - Formatting Objects for PDF3D. In M.N. Zuffo et al. (Eds.) *Proceedings of the 15th International Conference on 3D Web Technology. July 24 - 25, 2010 Los Angeles, CA, USA* (pp. 63-72). New York: ACM.

ECMA International. (2007). *Standard ECME-363, Universal 3d File Format (4th edition)*. Retrieved April 4, 2011 from http://www.ecma-international.org/publications/standards/Ecma-363.htm.

Edwards, S.J. (1994). *Portable Game Notation Specification and Implementation Guide.* Retrieved April 4, 2011 from http://www.chessclub.com/help/PGN-spec.

Evans, L. & Wiswell T. (1953). *Championship chess and checkers for all*. New York: A.S. Barnes & Co.

FIDE – World Chess Federation. (2008). *Actual Handbook E.1.01B. Appendices C. Algebraic notation.* Retrieved April 4, 2011 from http://www.fide.com/component/handbook/?id=125&view=article.

ISO 19005-1:2005. (2005). *Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1)*. ISO, Geneva, Switzerland.

ISO 32000-1:2008. (2008). *Document management – Portable document format – Part 1: PDF 1.7*. ISO, Geneva, Switzerland.

Khronos Group. (2011). *WebGL Specification Version 1.0*. Retrieved April 4, 2011 from https://www.khronos.org/registry/webgl/specs/1.0/.

W3C. (2006). *Extensible Stylesheet Language (XSL) Version 1*. Retrieved April 4, 2011 from http://www.w3.org/TR/xsl11/.

W3C. (2009). *Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) specification*. Retrieved April 4, 2011 from http://www.w3.org/TR/CSS2/.

W3C. (2011). *HTML 5*. Retrieved April 4, 2011 from http://www.w3.org/TR/html5/.

EPUB Use in Digital Libraries: Developing an Online EPUB Creator Application

Zeki Çelikbaş

İstanbul Technical University. 34469 Maslak, İstanbul, Turkey. Email: celikbas@itu.edu.tr

Abstract: *Libraries have a unique role as a bridge between users and information containers. Currently, information containers usually consist of books, journals and electronic files. While the format of information containers changes over the years, the library's role stays the same. As digital libraries proliferate, the classical library has a new role as information producer. In this paper, we discuss the EPUB format, developed by the International Digital Publishing Forum (IDPF) as an open standard for e-books, and its use as a new information container format in libraries and digital libraries. Several different formats (e.g., .doc, .odf, .txt, .html, .DjVu) are used in digital libraries, but the EPUB format stands out as it has open file structure and is easy to produce and use. Because of its ease of use, mobile devices and e-book readers such as Kindle immediately started to support it. As EPUB is an open and free format, it enables publishers and digital libraries to publish license-free e-books. We review the EPUB format in this paper and show how to produce an EPUB file along with its use with digital library applications such as DSpace, EPrints and Greenstone.*

Keywords: *Digital libraries; eBooks; EPUB.*

Introduction

This work is divided into two parts. First, we discuss the basic features of EPUB format for librarians on the basis of the literature and describe its structure. Second, we describe an EPUB application that we developed to create .epub files and use it in digital libraries and institutional repositories.

Why EPUB Format?

We have been working on two separate projects: Creating an institutional archive and developing an interlibrary loan (ILL) application. To accomplish both tasks, electronic documents should first be created. Currently, Adobe's portable document format (PDF) is used extensively as an e-book format. The EPUB document format as an open structure is being increasingly used because it has no license restrictions and is easy to use. We decided to use EPUB format for electronic documents to exchange theses and chapters of books among different Turkish institutions as we needed a quick and simple tool. The proliferation of and widespread use of electronic readers and the rise of EPUB format use played an important role in choosing the EPUB file format. Table 1 compares the two formats, PDF and EPUB.

Table 1. Comparison of PDF and EPUB

Product Format	Filename extension	Image support	Interactivity support	Reflowable content	Open standard	Bookmarking	DRM support
PDF	.pdf	Yes	Yes	No	Yes	Yes	Yes
EPUB	.EPUB	Yes	No (Yes*)	Yes	Yes	No	Yes

* Interactivity is supported in new version.
Source: http://www.anypdftools.com/ebook-reading/pdf-vs-epub.html

What is EPUB?

EPUB is a free and open standard for eBooks maintained by the International Digital Publishing Forum (IDPF) (EPUB 3 Overview, 2011). The EPUB file format was designed to be open to anyone to

use or to develop it and is reflowable (the text can be resized and rearranged to suit whatever display it is being read on).

The EPUB version 2.0.1 consists of three specifications:

- *Open Publication Structure* (OPS) contains the formatting of its contentX
- *Open Packaging Format* (OPF) describes the structure of the .epub file in XMLX uze
- *Open Container Format* (OCF) collects all files as a zip archive,

EPUB internally uses XHTML or DTBook (an XML standard provided by the DAISY Consortium) to represent the text and structure of the content of documents, and a subset of CSS to provide layout and formatting. XML is used to create the document manifest, table of contents, and EPUB metadata. Finally, the files are bundled in a zip file as a packaging format. For more information, see EPUB (2011) and EPUB 3 Overview (2011).

Table 2. A short list of EPUB viewers and software for EPUB files.

Software	**Platform**	**DRM formats supported**	**Notes**
Adobe Digital Editions	Windows, Mac OS X	Adobe Content Server	Requires online activation.
Aldiko	Android	Adobe Content Server	Supports EPUB for Android phones.
Bluefire Reader	iOS	Adobe Content Server	Supports EPUB and PDF.
Bookworm	Web	?	Free open source online EPUB reader.
Calibre	Windows, Mac OS X, GNU/Linux	None	More often used for library management, conversion, and transferring to devices than reading.
CoolReader	Windows, GNU/Linux, Android	None	XML/CSS based E-Book reader for desktops and handheld devices. Supported formats: FB2, TXT, RTF, TCR, HTML, EPUB, CHM.
EPUBReader	Firefox add-on	None	Enables reading EPUB-files from within Firefox.
FBReader	Windows, GNU/Linux, PDAs	?	Incomplete EPUB support.
FBReaderJ	Android	?	Open source.
Freda	Windows Mobile, Windows Phone 7	None	
Google Books	Web	?	Supports EPUB and PDF.
ICE Book Reader Professional	Windows	None	

Advantages of Using EPUB

As can be seen from the structure of EPUB files, the EPUB format has an open structure, is easy to create and can be integrated into long term preservation plans. For long term preservation, the text, metadata and multimedia content need to be converted successfully (Fig. 1). EPUB files can easily be exported to other file formats.

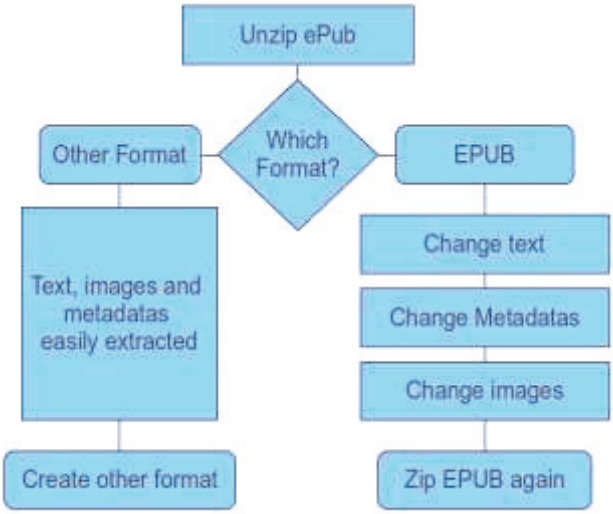

Figure 1. EPUB flow chart for long term preservation

How to Create an EPUB File

To create an EPUB file, one of the following methods can be used (Castro, 2011):

- Coding with any text editor;
- Using a special application (Calibre, Adobe InDesign, eCub, etc.); and
- Using conversion tools (ePub Tools, eScape - ePub Creator, etc.).

Aaron DeMott's tutorial on "How to Make an EPUB eBook by Hand" is a very informative document to understand how to create an EPUB file (DeMott, 2011). An EPUB file is created in three main steps:

1. Prepare the content;
2. Prepare the XML files (container.xml, book.opf, toc.ncx and all eBook content in one or more content.xml files); and
3. Put in the container.

The zip file layout should look something like this:

```
- mimetype
META-INF/
  - container.xml
OEBPS/
  images/
  - book.opf
  - toc.ncx
  - stylesheet.css
  - content.xhtml
```

EPUB Use in Libraries

Many libraries have started to lend e-books and e-book readers to their users. Most e-book readers support the EPUB format (see, for instance, http://wiki.mobileread.com/wiki/EBook_Lending_Libraries; http://sonysearch.overdrive.com/; http://reeed.ru/en_libraries_epub.php). All one needs to borrow an e-book from a library is an e-reader and a library card with an Adobe ID. Most libraries that lend e-books use the Adobe Digital Rights Management (DRM) protection system. This means that most dedicated e-readers (e.g., Sony,

Kobo, and so on) and several e-reader apps (e.g., Bluefire, OverDrive) can be used to read these DRM-protected library e-books (Reading, 2011).

Currently, the official EPUB standard does not include any specifications for DRM. However, most EPUB distributors at this time use the Adobe ADEPT DRM system, which means that EPUB files should be readable in most DRM-enabled devices (excluding Apple's iPad) (What is ePub, 2011).

Because of the XML standard, it is very easy to integrate DRM into EPUB. An EPUB file can optionally contain DRM , but it is not required by the specifications. Today, many books are sold with EPUB-integrated DRM support.

Adobe's DRM is a proprietary system, but Open Digital Rights Language (ODRL) as an open standard for rights management is on its way. The ODRL Initiative is an international effort aimed at developing and promoting an open standard for policy expressions. ODRL provides flexible and interoperable mechanisms to support transparent and innovative use of digital content in publishing, distribution and consumption of digital media across all sectors and communities (ODRL, 2011).

EPUB Application Development

There have been special applications for the interlibrary electronic document exchange for a long time. ILLiad, Ariel and OCLC's applications and even librarians' corporate email addresses have been used for document exchange. Librarians used various file formats for this purpose such as multipage TIFF, PDF, original office document files and image files. When we developed KITS, an ILL module used by many university libraries in Turkey, we had to tackle the issue of exchanging different ILL transaction files among libraries. We chose the EPUB open standard because of its simple structure and developed the Online EPUB Creator.

The Online EPUB Creator

This application has been created to produce a file from scanned images for librarians who will be exchanging electronic documents among different libraries.

The Online EPUB Creator runs on a web server. The short description of this application is that it produces an EPUB file from the images that are transferred by means of web forms. A working copy of the Online EPUB Creator was installed on a public server. Another version is still serving KITS members. Both applications have their own features. The following web addresses can be used to access the application: http://epub.ankos.gen.tr/ and http://kits.ankos.gen.tr/onlinepub.

The files generated are held on the system for a certain period of time (24 hours to 15 days). At the end of this period, files are automatically deleted, although users who have the delete link activated can delete files any time they wish. This sounds a bit like a public file sharing system, but there are several public document creation services working in the same manner.

Application Infrastructure

The application we developed consists of the following components:

- Web server (Apache);
- PHP scripting language;
- Asbjorn Grandt's zip php class (Grandt, 2011); File download class (A php class sends the requested files to the clients and prevents unwanted downloads) (Ashraf, 2011); and
- EPUBCheck (to confirm compliance of generated EPUB with standards described by the International Digital Publishing Forum (IDPF)) (EPUBCheck, 2011).

An open source Java tool, called EPUBCheck, validates and detects errors in the structural markup (OPS, OPF, and OCF) as well as the XHTML and image files. The tool can be run from the command line, or used in webapps and applications as a library. Fig. 2 shows how this check can be made on any EPUB file. Thus, no structure or file errors are encountered in an EPUB file. This file can then be used as an e-book.

```
$ java -jar epubcheck-1.2.jar
uploads/49a5e86d5cde/ebook.epub

Epubcheck Version 1.2


No errors or warnings detected
```

Figure 2. EPUBCheck validation and error detection tool

Creating EPUB files using our application is very easy. All one needs to do is to enter images into the form elements in the web page and then press the "Create Epub" button. Figures 3 and 4 show the input and confirmation screens. The latter shows the URL link of the newly-created EPUB file along with the delete link.

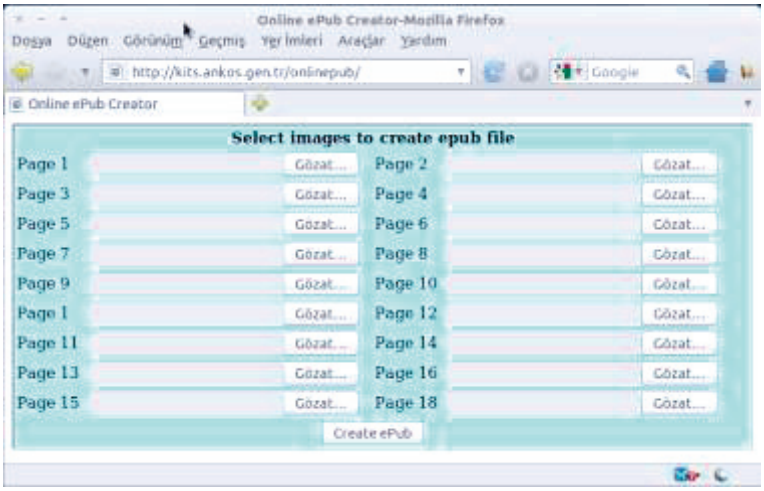

Figure 3. Online EPUB Creator application

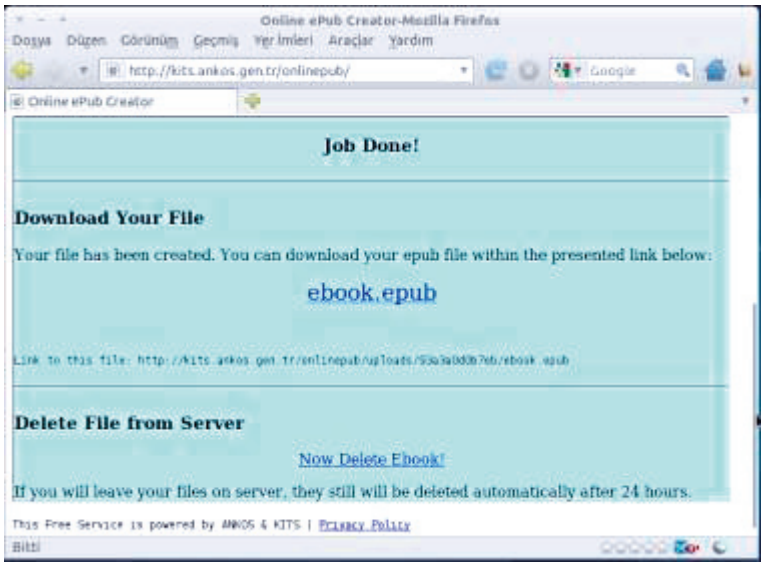

Figure 4. Online EPUB Creator's confirmation screen

All the processed files are located under a randomly created directory. The "ebook.epub" created with the name of the electronic document is also available. The OpenURL link for the distribution of this information is also created. At the bottom of Figure 4 a special link contains information on deleting this file. When this link is clicked, the created directory and all of its contents get deleted

permanently. A link to a license file containing terms of use of this service is also given. While this application is concentrated on creating public EPUB files, the ILL application (KITS) has more restrictions such as the number of times a specific EPUB file can be downloaded before it gets deleted. Restrictions can be defined by the creator. The next version of the system will include more features such as the ability to define the number of times image files can be replicated, requiring more complex Javascript libraries and AJAX capabilities.

EPUB and Digital Libraries

The most important factor in serving EPUB files with digital library applications is that of compatibility with librarianship standards. When selecting a digital library application to be used in an institution, the following three criteria should be considered:

1. EPUB support;
2. Compliance with standards; and
3. Ease of installation and usage.

Three digital library applications (DSpace, EPrints, and Greenstone) were tested for this purpose. As EPUB is quite a new standard, these three applications do not support EPUB files, although they are compatible with standards. All three are easy to use although DSpace and EPrints are harder to implement. Table 3 lists major characteristics of these digital applications.

Table 3. Characteristics of digital library applications

	Bulk Data Input	Bulk Metadata Input	Browsing and Listing	Plug-in Support for EPUB Files
DSpace	No	No	Yes	No
EPrints	No	No	Yes	No
Greenstone	Yes	Yes	Yes	No

As Greenstone offers the opportunity to add bulk file and metadata input, we decided to use it in our project. None of them has plug-in support for EPUB files, although we plan to develop one for Greenstone. We configured Greenstone to process EPUB files using an unknown plug-in (Fig. 5). We added an EPUB mimetype to the Dublin Core format to run this plug-in.

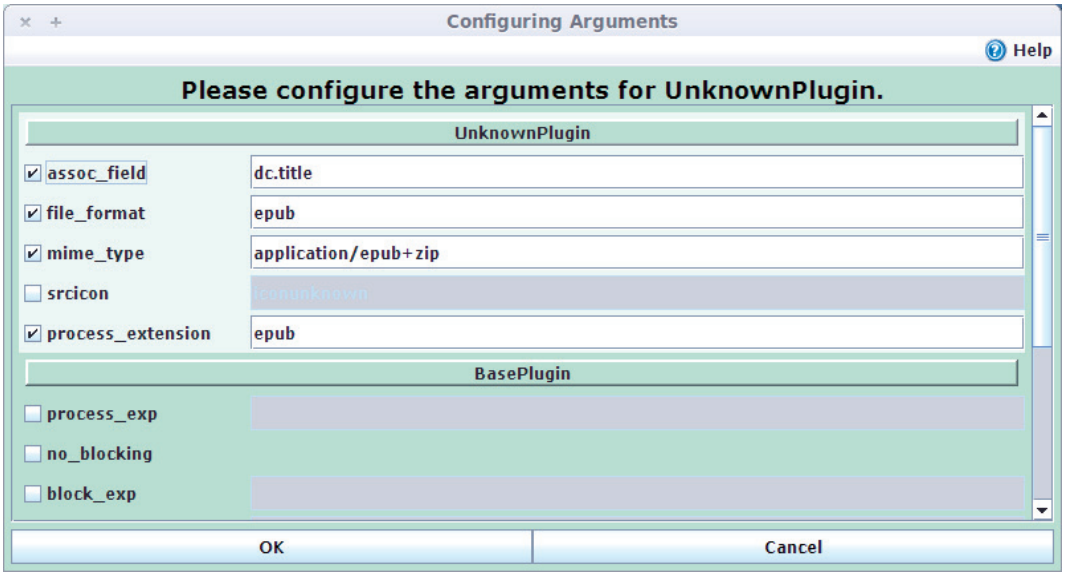

Figure 5. EPUB mimetype

Digital library applications do not yet have plug-in support to display EPUB files. EPUB will benefit from such third party applications. EPUB display applications were listed earlier in this paper in the "What is Epub?" section. We selected EPUBReader (http://www.epubread.com/) as our EPUB viewer. This application works as a Firefox browser plug-in, which means that the application we developed will be accessible through Firefox and offers platform-independent access to users. Needless to say, those wishing to see EPUB files using a different viewer can do so.

Conclusion

Although EPUB is not a new standard, it has not been fully supported by libraries and digital library applications for several reasons. Among them are:

- EPUB creation tools are insufficient;
- Adequate tools do not exist for Digital Rights Management; and
- Publications on EPUB are scarce.

Some libraries have started to buy e-book readers for their users, yet not all of them have EPUB support. Librarians must consider having this feature available in their e-readers.

The Open Digital Rights Language (ODRL) is an XML-based language for expressing, providing and agreeing to rights information of any kind of content and can be easily integrated into EPUB files, thereby increasing their potential use in the near future.

References

Ashraf, Q. (2011). PHP File Download Class. *PHP Classes Repository - PHP Classes*. Retrieved, 4 April 2011, from http://www.phpclasses.org/package/6687-PHP-Serve-files-for-download.html.

Castro, E. (2010). EPUB straight to the point: Creating ebooks for the Apple iPad and other ereaders. Berkeley, CA: Peachpit Press, 2011. Retrieved, 14 May 2011, from http://www.ipdfconverter.com/blog/iphone-ipad/epub-straight-to-the-point-creating-ebooks-for-the-apple-ipad-and-other-ereaders-free-pdf-ebook-download.html.

DeMott, A. (2011). epub eBooks Tutorial. Retrieved, 4 April 2011, from http://www.jedisaber.com/ebooks/tutorial.asp.

EPUB - *Wikipedia, the free encyclopedia*. Retrieved, 03 April 2011, from http://en.wikipedia.org/wiki/EPUB.

EPUB 3 Overview. (2011, February 15). International Digital Publishing Forum | Trade and Standards Organization for the Digital Publishing Industry. Retrieved, 03 April 2011, from http://idpf.org/epub/30/spec/epub30-overview.html.

EPUBCheck - Validation tool for EPUB - Google Project Hosting. (2011). *Google Code*. Retrieved, 4 April 2011, from http://code.google.com/p/epubcheck.

Grandt, A (2011). Zip (archive, create zip file, zip class). *PHP Classes Repository - PHP Classes*. Retrieved, 4 April 2011, from http://www.phpclasses.org/package/6110-PHP-Create-archives-of-compressed-files-in-ZIP-format.html.

ODRL Initiative. (2011). Retrieved, 4 April 2011, from http://odrl.net/.

Reading Library eBooks on your iPad, Android, eReader and iPhone - epubBlog: News and information on the EPUB Standard. (2011). *Free EPUB eBooks for your iPad, Android, Kobo, Nook and Sony eReaders | Unleash Your Books*. Retrieved, 4 April 2011, from http://www.epubbooks.com/blog/668/library-books-on-your-ereader.

What is ePub? (2011). epubBlog: News and information on the EPUB Standard." *Free EPUB eBooks for your iPad, Android, Kobo, Nook and Sony eReaders | Unleash Your Books*. Retrieved, 4 April 2011, from http://www.epubbooks.com/blog/what-is-epub.

Hiperkitap: A Turkish E-book Database

Serdar Katipoğlu
Hiperlink Ltd., Mumhane Cad., Murakıp Sok., No: 13 Tophane, Istanbul, Turkey.
Email: serdar@hiperlink.com.tr

Abstract: *The purpose of this paper is to describe Hiperkitap, the first and only Turkish e-book database in Turkey. The paper starts with a brief overview of the developmental stages of the e-book database under the local market conditions in Turkey. This developmental period includes the establishment of relationships with publishers, authors, the publishers' association of Turkey, computer companies, information designers and librarians. Content provision under Turkish and E.U. copyright laws is a much debated issue in Turkey. Some 7500 books in Turkish from 220 Turkish publishing houses make this database attractive for local customers and also for foreign competitors. Rich content and up-to-date technology are long term strategic targets for this database. Mobile applications, mostly smart telephone applications, are determining the development policy of Hiperkitap today. This paper aims to be useful for understanding the social history of digital publishing and the e-book market in Turkey.*

Keywords: *Hiperkitap; ebooks; Turkish e-book database.*

Introduction

Sales people in Europe and the USA started to visit the academic market to present a new product, CD-ROMs, in the mid-1980s. At that time, CD-ROMs contained abstracts of journal articles. In the mid-1990s CDs contained encyclopedias, dictionaries and full-text journal articles. Librarians and academic staff found themselves split between the print resource lovers and electronic resource lovers. Sales people, and some librarians, had to overcome the resistance of academic staff to e-resources.

The academic Internet infrastructure in Turkey was nearing completion by the Turkish Academic Network and Information Center (ULAKBIM) at the end of the 1990s (ULAKBIM, 2010). This was a very big chance for e-resource companies. There was a highway but no cars on it. They used this chance. Turkish university libraries were targeted by the foreign database aggregators or single companies. They started to organize seminars, personal visits and on-site training for librarians. It was the introduction and promotion period. They were creating the database market and, furthermore, identifying and attracting potential users. There was one interesting issue, though: all databases were in English whereas the medium of instruction is Turkish at most of the universities in Turkey. Exception to this was "Index 95", the first Turkish popular journal database with scanned full-text articles released in 1995 in Istanbul, although it did not last very long and disappeared from the market. Alt Kitap Online (www.altkitap.com) was the first online publisher in Turkey; it is still active but with a low profile.

Consortium

As we know, the increasing number of databases on the world market brings a new type of organization for university libraries. This is the University Libraries Consortium. The appearance of academic databases in the Turkish market and the sharply increasing number of state and foundation universities in Turkey were good reasons to establish a consortium. The Anatolian University Libraries Consortium (ANKOS) was established in 2000 as the first consortium in Turkey with 12 member libraries and three database subscriptions. Faculty members and students at universities all over Turkey as well as the staff of some research institutions now have access to bibliographic and full-text databases. ANKOS holds annual meetings attended by all ANKOS librarians and database companies. The former chair of the Steering Committee of ANKOS and a colleague suggested a supplementary mission for the consortium: "In the light of the successes of the

first 10 years of ANKOS, we believe that it should continue to develop its role in providing professional training and developing skills, in addition to acquiring electronic resources" (Erdogan & Karasozen, 2009).

ULAKBIM, on the other hand, created Electronic Resources National Academic License (EKUAL), signed agreements with database companies and then opened these databases to libraries free of charge. It has offered 20 full-text and bibliographic academic databases since March 2008. EKUAL pays for national site licenses. The significance of EKUAL lies in its structure. It is not a consortium but an independent state organization. ULAKBIM had started to design and produce national databases in various scientific fields in the 1990s. These were the Turkish Medical Database, Turkish Life Sciences Database, Turkish Social Sciences Database, Turkish Engineering and Basic Sciences Database. These databases contained articles published in different languages in Turkish scholarly journals.

Associations

The Turkish Librarians' Association was founded in 1949 and the University and Research Libraries Association (UNAK) was established in 1991. Every year the last week of March is observed as library week in Turkey. ANKOS, ULAKBIM and the library associations are active leaders in the librarianship and information area.

The rapid growth of the database market together with ANKOS, EKUAL and association activities, events, professional training seminars and conferences have developed the skills and attitudes of librarians and academic staff. As a result, this social transformation has created an e-resource culture among these professional and academic groups.

On another front, the wide use of databases brings academic success. ANKOS made a large contribution to increasing the scientific publication rate by Turkish scientists at the international level (Karasozen, 2008).

Digital Culture

The highly accepted e-resource culture, and the predominance of databases in the English language were important factors in creating demand for a Turkish e-book database.

Social, academic and professional demands for a Turkish database were heard at library conferences, in library visits and among librarians. Local vendors and publishing companies were being pressured for a Turkish e-book database by these potential users. This was wind of social change for the market

Electronic publishing was also a very hot topic for the publishers in Turkey. The Turkish Publishers Association's 4th Conference was held on 6-7 May 2010. The conference theme was "Legal and Financial Aspects of Digital Publishing". The final announcement emphasized that the E.U. Information Society criteria should be followed for digital publishing. They suggested that those very clear articles should be included in the Intellectual Property Law. They recommended that the Ministry of National Education be the organizer and supporter for digital publishing (Turkish Publishers Association, 2011).

Value Added Tax (VAT) is 8% for print books and 18% for e-software and hardware items in Turkey. In terms of VAT for e-books, there is no clear official decision whether it should be 8% or 18%. The Turkish Publishers Association has suggested to the Ministry of Culture that VAT for e-books should be the same as for print books: 8%.

The Turkish Publishers Association is very supportive of electronic publishing within the context of business ethics and the Intellectual Property Law.

Hiperlink

Hiperlink was established in 2006 in the social context described above. Hiperlink is a book and non-book materials supplier in Turkey. It establishes and designs libraries and collections acting as a kind of professional institute for librarianship.

It signed agreements with two global companies (LexisNexis and Knovel) in 2008, its first international dealing. The nature of the company was in transition; it is becoming global. Additionally, it had started to publish print books in 2007.

Hiperlink hired two staff members and a temporary working group which started to discuss and investigate the "Hiperkitap" concept as a project in 2007. The search for software and an e-book reader platform followed in 2008.

They started to visit publishers for content in 2009. This was a pessimistic period. The publishers were not ready to accept digital publishing. They were publishing print books only. Their copyright agreements with authors and agents contained only articles for print editions. A great deal of time was spent to persuade the publishers to complete agreements for online copies. These were "warm-up laps".

They first signed agreements with 100 publishers for 3500 online editions of books. All these books were prepared digitally for the press. All of them were digital-born in a sense. Apart from the online editions, the publishers were still giving the print editions of books in large numbers to Hiperlink. This was another digitizing process. Maybe they were afraid of this transition period or they were not ready psychologically to be involved and take responsibility in the frame of the copyright law. So far, the copyright of all books translated into Turkish from other languages is only for the print edition.

Hiperkitap as a database was accepted and socially approved by the ANKOS administration in November of 2009. Under the ANKOS consortium 24 university libraries subscribed to Hiperkitap that year. This number has increased to 36 subscriptions by the end of 2010, 34 of them university libraries and the remaining two public libraries. One of the public libraries is strictly a digital library. It has only Hiperkitap as a digital collection, no other print or digital books (Esenler, 2010).

Nowadays all problems seem to have been solved or are about to be. Today, Hiperkitap has obtained 7500 books from 220 different local publishers (Table 1). The collection consists of academic and general culture subjects and is appropriate for university and public libraries in terms of content. As a result 55 university libraries, two public libraries and two other higher education institutes are subscribed to Hiperkitap in 2011 (Table 2).

Table 1. Publishers and books

Year	**2009**	**2010**	**2011**
Publishers	115	205	220
Books	3550	7000	7500

Table 2. Subscribers

Year	**2010**	**2011**
Universities	34	55
Public libraries	2	2
Other institutions	-	2
Total	**36**	**59**

System Technical Brief

The system consists of two physical dedicated servers (Windows Server 2003 x64 architecture) and separate IPs. The Apache web server redirects Java requests to eight load-balancing tomcat nodes residing on both servers. The project makes use of robust open source Java libraries such as log4j, lucene and struts, while a PHP based Content Management System is used for the front end; a critical business logic/layer is coded in J2EE. Thanks to caching mechanisms business objects such as

Book, Page or Reader are represented in memory. The scanning-indexing-searching process is completely flexible and scalable.

System Management Brief

A book tagging process allows the editors to choose books in order to offer clients/readers different book collections and avoid/filter undesired content. By using IP filters and client domain name definitions, client browsers are instantly recognized and offered only the appropriate content/search results.

Book pages are served as ordinary PDF files, thus allowing a huge range of browsers to access the book content without installing messy Java/Flash based plugins or readers. Simplicity of pages allow any mobile device with a pdf viewer to access the book content easily. Page navigation, value added services and related content on the WWW can also be embedded in books.

Functionality and Accessibility

The search engine offers basic and advanced options and a "smart" option as a beta. Keyword search within full text is functional. All browsers are compatible with the system. For the best performance, Microsoft Internet Explorer 6.0 and more recent editions and Adobe Reader 9.0 are recommended. My Shelf is available as a module.

MARC records are searchable in the OPACs of library automation systems. MARC records can be given to some aggregators' search engines in advance (for instance, Discovery, AtoZ and Serial Solutions). Any user can access the database via a mobile or smart phone in the IP-based area. Hiperkitap staff supply extra technical assistance and permanent social support to the subscribers. Some 350,000 pages of e-books were viewed through Hiperkitap in 2010 and the figure is expected to increase considerably in 2011.

Table 3. Functionality and accessibility

Format of e-books	Access	Method of access	Citations
PDF	IP-based	Annual subscription	APA and MLA

Conclusions

This is the first time in Turkey that a Turkish e-book database has been created. Mobile applications, mostly smart telephone applications, are determining the development policy of Hiperkitap. The strategic aim is to provide access to up-to-date e-books using technology at the global level.

References

Erdogan, P. & Karasozen, B. (2009) Portrait of a consortium: ANKOS (Anatolian University Libraries Consortium). *Journal of Academic Librarianship,* 35(4), 377-385.

Esenler. (2010). Esenler Dijital Kutuphanesi. Retrieved May 2, 2011 from http://www.esenler.bel.tr/index.php?c=haberler&bid=847

Karasozen, B. (2008). Consortial usage of electronic journals in Turkey. *LIBER Quarterly*, 18(3/4), 464-469.

Turkish Publishers Association. (2011). 4th Conference Final Announcement 6-7 May 2010, Retrieved May 2, 2011 from http://www.turkyaybir.org.tr/.

ULAKBIM. (2010). ULAKNET. Retrieved May 2, 2011 from http://www.ulakbim.gov.tr/hakkimizda/tarihce/ulaknet/dunbugun.uhtml.

Information Workers and their Personal Information Management: a Literature Review

A.A.J. (Jos) van Helvoort

The Hague University of Applied Sciences, Academy of ICT&Media, Post Box 13336, 2501 EH The Hague, The Netherlands. Email: a.a.j.vanhelvoort@hhs.nl

Abstract: *The research described in this paper provides insights into tools and methods which are used by professional information workers to keep and to manage their personal information. A literature study was carried out on 23 scholarly papers and articles, retrieved from the ACM Digital Library and Library and Information Science Abstracts (LISA). The research questions were:*

- *How do information workers keep and manage their information sources?*
- *What aims do they have when building personal information collections?*
- *What problems do they experience with the use and management of their personal collections?*

The main conclusion from the literature is that professional information workers use different tools and approaches for personal information management, depending on their personal style, the types of information in their collections and the devices which they use for retrieval. The main problem that they experience is that of information fragmentation over different collections and different devices. These findings can provide input for improvement of information literacy curricula in Higher Education. It has been remarked that scholarly research and literature on Personal Information Management do not pay a lot of attention to the keeping and management of (bibliographic) data from external documentation. How people process the information from those sources and how this stimulates their personal learning, is completely overlooked.

Keywords: *Personal Information Management; information literacy; higher education.*

Introduction

The research reported in this paper has been conducted in the context of my research activities in Information Literacy. The ACRL Information Literacy Competency Standards for Higher Education mention as performance indicator 2.5: "the information literate student extracts, records, and manages the information and its sources" (ACRL, 2000). The aim of the research presented here is to construct an overview of tools and methods which are used by professional information workers to keep and to manage their information. This overview might suggest directions for the improvement of Information Literacy Curricula in Higher Education (Jones 2007, p. 457).

Research Approach

Research Questions

The main question of the research was: What kind of tools and approaches do information workers use to keep and to manage their information sources?

Subsequent questions were on the aims they have when keeping information and the problems they experience.

These research questions belong to the field of Personal Information Management (PIM), a research field that includes users' activities to acquire or to create information items and the retrieval, use and distribution of them (Jones 2007, p. 457). However, in the context of the ACRL performance indicator 2.5 the present research is focussed on storing, organising and maintaining information with the aim of creating possibilities to locate it for later use.

Methods

The research was conducted as a literature study. The first step was desk research of relevant scholarly literature. The second step was the analysis of its content. For the content analysis the results of the desk research were grouped in themes (for instance "Types of collections that people create") and for each theme the main findings are presented.

Search Strategy

The first document that was read on the topic PIM was Jones' chapter in the Annual Review of Information Science and Technology (2007). In the next step of the orientation with Google Scholar it was observed that a lot of scholarly literature on PIM consists of conference contributions from which the papers are published in the ACM Digital Library. A search in the ACM Digital Library with the *discovered terms* "Personal Information Management" and "User Studies" resulted in a hit list with fourteen references (27 November 2010). After further reading, one of them appeared not to be relevant to the research question. Ten of the thirteen relevant documents were full text available and were used for this study.

Because the ACM Digital Library focuses on Computer Science the search was extended with a query in the main source for Library and Information Science LISA on 28 November. Search terms were "personal information management" AND NOT "book review", limited to peer reviewed journals in the English language. This query resulted in twenty one hits. After first reading it appeared that sixteen of them were relevant for my own research question. Thirteen titles were full text available and were used for the research. There was no overlap with the titles from the ACM Digital Library.

Findings

Types of Collections that People Create

The concept of PIM refers to information management activities in various domains: email messages, contact lists, files and documents in a "my documents" folder, references to web pages and tangible documents like photocopies, prints, handwritten notes and books. Jones refers to those domains as "Personal Information Collections" which together form ones "Personal Space of Information" (2007, p. 462). The elements of a collection mostly share a "particular technological format and are accessed through a particular application" (Boardman 2004, cited by Jones 2007, p. 462). Most collections described in the literature are local files, email and bookmarks. Remarkable is that in the Computer Science literature no attention is paid to the use of bibliographic software (for instance Endnote, Refworks or Zotero). From the titles which were retrieved from LISA only McGeachin (2004) and Yu (2008) pay explicit attention to bibliographic citation management by scholars.

How People Organise their Personal Information Collections

Indratmo and Vassileva (2008) distinguish five different approaches as to how people organise their personal information: hierarchical, flat, linear, spatial and networked. Besides these five approaches a sixth approach can be distinguished which uses a minimum of organisation and is for that reason very popular to end users of information: full text search. All these six approaches will be discussed here in more detail while remarks and insights from different authors will be added to the elaboration.

Hierarchical approach

The hierarchical organisation of information items (for instance in folders and subfolders) is most commonly used by information workers. People are very familiar with it and therefore the process of organising information in a hierarchical way is very intuitive (Indratmo & Vassileva 2008, p. 2). The drawbacks are the problems in creating a consistent structure and difficulties in choosing which category one should file an item (Bergman et al. 2008[a], p. 235; Bergman et al. 2008[b], p. 3; Indratmo & Vassileva 2008, p. 3; Jones 2007, p. 473). Nevertheless, the hierarchical approach seems to be a good solution if the information user needs a complete overview of all the relevant documents (high recall).

Indratmo and Vassileva refer to this situation in the context of project management (2008, p. 10). In practice the use of hierarchical folders is quite common for information that belongs to specific tasks when it is easy to determine to which task a document belongs (one folder for each task). However, the hierarchical approach hinders retrieval by means of mobile devices such as smartphones, due to their small screens. For those devices, full text search or sending relevant documents to the working space seem to be better solutions (Glória et al. 2010).

Flat approach

With the "flat approach" Indratmo and Vassileva refer to the practice of labelling or tagging information items (Indratmo & Vassileva 2008, p. 4 ff.). The most striking advantage of the flat organisation method is that users can classify an information item into multiple categories simply by assigning multiple tags to the item (Ibid. p. 4). The drawback is of course the risk of "noise" in the retrieval phase. Indratmo and Vassileva therefore emphasise that the flat approach is particularly suitable for more general collections of information (2008, p. 10). The flat approach is best known in different tools for online bookmarking (for instance Delicious).

Linear approach

With the linear approach of storing and retrieving information people use the function of an information system to sort the output of a query or an information collection, for instance chronologically or alphabetically. This approach is well known in email clients,where one can sort messages under the sender ("from") or a date ("received") or in applications to browse photos ("date created"). It is often used as a second approach within a folder or a set of information items with a common tag (Boardman & Sasse 2004, p. 587). This technique is also meant by Indratmo and Vassileva when they plead for providing multiple visualisations (Indratmo & Vassileva 2008, p. 11).

The common habit of storing books or prints of journal articles alphabetically, under author names, is another example of the linear approach.

Spatial approach

A spatial approach to organising information items means that frequently or recently used items are presented for quick access on, for instance, the user's computer desktop. Indratmo and Vassileva describe a tool for personal information management that helps users to arrange their information items on a special desktop without using folders (TMC, Time-Machine Computing. Indratmo & Vassileva 2008, p. 7) but more frequently this approach will be used by information workers in their document applications with a sidebar that allows them to find a shortcut to the last used documents. The Inbox of an email client is also often used as a spatial tool in which the current work can be found.

The spatial approach is often combined with "past interaction" (Bergman et al. 2003, p. 875-876; Indratmo & Vassileva 2008, p. 12). The electronic traces that people leave on the internet and on their local systems are, in this approach, used for quick access to their documents. The best known example of this approach is without doubt the history button of a web browser. "Past interaction" is also the principle that is used for access to recently used documents in text editor applications, as discussed earlier.

A spatial approach could also be a solution for archiving purposes. Keeping apart the most valuable digital information makes it easier to manage it for availability in the long term (Cushing 2010, pp. 305-307). However, it appears that archiving of personal information still lacks enough attention in the research and literature on PIM (Ibid.; Williams et al. 2009, p. 340).

Network approach

This approach is a rather sophisticated method of organising information. It works with the linking principle that is also used to organise the World Wide Web. Creating links between information items (even between items in different information collections) is quite easy but in the retrieval phase it is hard to navigate and to find your way in the network (Indratmo & Vassileva 2008, p. 8). Just as on

the World Wide Web it is easy to find one relevant item but finding all relevant items ("high recall") is a complicated task.

Minimal structure: full text search

With full text search in their personal information collections people use characteristic phrases or author names to refind the information items. Search technology potentially solves the problems that information users have with organising their information items in hierarchical folders (Bergman et al. 2008[b], pp. 3-4). However, empirical studies show that, for personal information management, people still prefer a hierarchical folder and navigation method over search (Barreau 2008, p. 312; Bergman et al. 2008[b], p. 16). Bergman et al. suggest that retrieval behaviour in the context of personal information management differs from retrieval behaviour on the World Wide Web because "in PIM, the same user both organizes and retrieves the information" (Ibid. p. 20). Moreover, they argue that the organisation of information items in folders functions as "sense-making" (see also Jones 2007, p. 487).

For full text search, the role of memory is essential. People have to remember characteristic words from the text, title or other bibliographic data. Sometimes this is a problem. A lot of the conference literature is on experiments to improve integrated desktop search with autobiographic information that people remember (Dumais et al. 2003; Elsweiler et al. 2007; Gonçalves & Jorge 2008) or associations with events and persons (Chau et al. 2008). These kinds of tools make it possible to get an answer to a query like: "the document that I received from the woman whom I met at the conference in Budapest". Bergman et al. confirm the importance of this kind of "subjective attributes" of information items in personal information management (Bergman et al. 2003, p. 873; Bergman et al 2008a, p. 236). However, Elsweiler, Baillie and Ruthven found that, at least for email, this kind of information is less often remembered than the topic of the message or the reason it was sent (Elsweiler et al. 2008, p. 31). In a more recent research article they concluded that this might be true but that it does not mean that remembering the topics of email messages also leads to better performance of refinding tasks. They conclude that finding tools in email clients should provide better support for semantic retrieval of information (Elsweiler et al. 2009, p. 24).

Different approaches combined

Boardman and Sasse discovered that people use different organising strategies for different collections. More effort is invested in organising their document files than in organising email or bookmarks because of the effort they have initially spent on creating their documents. Email collections are typically retrieved with the linear approach (Boardman & Sasse 2004, p. 589) and particularly "no filers" have a good memory for the information in their email collection (Elsweiler et al. 2008, p. 22). More in general the authors of the previously mentioned article plead for flexibility in refinding tools when designing interfaces for email clients. It appears that for the retrieval of information most people prefer a mix of different tools and approaches (Elsweiler et al. 2008; but also Elsweiler et al. 2007, p. 934). An example of this is a hierarchical search combined with alphabetical or rather chronological presentation of the results (for instance: recently used documents on the top of the list).

Descriptions that People Use to Keep their Individual Information Items

Boardman and Sasse (2004, p. 586) classified folder names that people used for the folders with their email, files and bookmarks. They found names for projects (short term activities), document properties (for instance 'letters'), roles (long term activities, for instance 'teaching'), contacts and topics/interests. In their research on the way in which people organise their information items, Bergman et al. 2008[a] (p. 239) found that most computer users tend to use project names for their folders more than file properties (for instance their technical 'format'). This is consistent with Kwasnik's findings (1989, cited by Williams et al. 2009, p. 352). However, it is contrary to Barreau's conclusion that for electronic documents the managers she interviewed most often used the document attribute "Form" (which includes types of documents as well as file formats) for folder names. But it has to be said that she only interviewed four managers and that those managers recognised that

"what is easy and convenient for them may not be best practice for their organization as a whole" (Barreau 2008, p. 316).

Aims that People have when Building Personal Information Collections

During the performance of a task people seek information to address an immediate information need. They keep it in their systems, for instance in their bookmark collection or in their email client as a self-addressed email message for future use (Bruce 2005). However, the collecting of information is not always purposeful. Coughlan and Johnson found how creative practitioners collected interesting materials and ideas which came along more or less accidentally and which were not really related to the projects they were currently working on (Coughlan & Johnson 2009, pp. 7-9).

In daily life people also encounter information incidentally (for instance by listening to a radio broadcast or talking with colleagues) and they match it to "anticipated needs" (Bruce 2005; Jones 2007, pp. 472-473). That is why Bruce (2005) remarked that enhancing "sensitivity and appropriate responses to *personal, anticipated information need* is […] a key component of information literacy".

"Keeping decisions" are fundamental to personal information management (Jones 2004). Keeping too many items may distract a person's attention from information that really is useful. On the other hand, the strategy of keeping nothing and trusting that one will retrieve it again on, for instance, the World Wide Web at the moment that one needs it, might lead to the situation that one forgets its existence entirely (Ibid.). Nevertheless, Barreau also found that managers trust more in searching again than in keeping bookmarks for information on the Web (Barreau 2008, p. 316).

The management of personal information collections is also a component of someone's learning process (Bergman et al. 2003, p. 872). Information items are read, interpreted and cognitively processed. As mentioned earlier, some authors believe that the majority of information users still prefer organising their information items in a hierarchical folder system because the classification of the items also functions as "sense making" (for instance Bergman et al. 2008[b], p. 20).

A third reason why people store information items is to remind them of a task that they have to perform (Boardman & Sasse 2004, p. 583; Elsweiler et al. 2007, p. 925). Particularly the email inbox is often used as a "to do" list (Barreau 2008, p. 310).

Yu (2008) noticed a remarkable alternative possible aim of gathering information by publishing scholars. The model that he proposes not only encourages scholars to create a personal information collection for citation purposes but also stimulates them to share their knowledge sets with other scholars. For the sharing scholar this has the advantage that it increases his visibility in the scholar world.

Problems Experienced

Fragmentation of information

According to Jones (2007, pp. 453-456) one of the biggest problems that people experience with personal information management is the fragmentation of information. As described earlier, information items are stored in a variety of information collections. Indratmo and Vassileva (2008, p. 10) give the example of information items related to a project but that are stored in a file system, an email client and a calendar. This enumeration can easily be expanded with, for instance, financial information from an administrative system and bibliographic records from a citation manager.

Bergman et al. propose a couple of improvements on interface design to enable computer users to organise their information items in a project based way (Bergman et al. 2008[a], pp. 241-243). However, the most commonly used approach to integrate information items from different collections is without doubt "integration through search" with "cross-format" desktop search facilities, for instance Google Desktop, or Spotlight in the Mac operating system. But this technological approach is not a solution for all information management problems, for instance not for the problem of multiple versions of documents (Jones 2007, p. 487). Moreover, as discussed earlier, search engine technology has the disadvantage that it presents access to individual documents and not to sets of documents that relate to each other.

Fragmentation of information not only occurs across different information collections but more and more also across various devices such as desktops, laptops, smart phones and personal digital assistants (Jones 2007, p. 475). Storing information items "in the cloud" could contribute to a solution for this different devices problem, although Dearman and Pierce report that academic professionals are more willing to use such web-based services than employees in industry (2008, p. 773).

People have to realise that they have an information need

Jones (2007, p. 468) emphasises that there is an important step preceding the retrieval of information from the personal space of information. He calls this "Remembering To Look". This step is comparable to "Recognising an Information Need". When people do not realise that they have an information need for the task at hand they will indeed forget to look up the information in their information space. Indratmo and Vassileva discuss a solution for this problem by the use of "future" or "planned interaction" records that remind information users of "potentially relevant documents" (Indratmo & Vassileva 2008, p. 12; but also Chernov 2008). Jones, however, also made a remarkable statement when he observed that keeping information also enables people to remember to look for it at a later time when the need arises (Jones 2007, p. 477). He even argues that putting an information item in a folder enables the cognitive processing of the content of the item (Ibid. p. 487). However, this is in contrast with Elsweiler and others' earlier mentioned conclusion that "no filers" have a good memory for the information in their email collections.

Keeping decisions

In the previous section I discussed the problems that information users often have with "keeping decisions" (Jones 2004). To solve the keeping problem Bergman et al. suggest a sorting tool for information retrieval that puts less used items at the bottom of the result list with a different appearance (for instance a smaller font and in a grey background) (Bergman et al. 2008[a], p. 243). This solution could also be used to present older versions of a document. As Bergman et al. argue, these design solutions combine the advantages of the "keep" option (the information item may be accessed in case it is unexpectedly needed in the future) with the advantages of the "delete" option (the information item no longer distracts the user's attention when he is looking for relevant information).

Discussion and Conclusions

The literature research described in this contribution was restricted to scholarly papers and articles from two different sources: the ACM Digital Library and LISA. This could be regarded as a limitation of the research but I do not have the impression that trends or developments in PIM are missed.

The research was started with the expectation that findings could contribute to the improvement of Information Literacy Education. In that context the results were a bit disappointing. Most of the literature that has been studied was on the management of files and emails. Only two articles discussed the management of (scholarly) literature and no attention at all was given to the way in which professionals extract that information. Only Williams et al. (2009, p. 350) remarked that this cognitive activity is indeed largely overlooked. The different contributions of Bergman et al. to PIM research confirm that the management of personal information stimulates one's learning process but it is not made very clear how this happens. This leads me to the conclusion that cognitive processing of information deserves more attention in PIM research. Potential research questions for this are:

- Which tools (for instance local bookmarks, online bookmarks or personal notes) do people use to manage references to information items that they have encountered and evaluate as possibly useful for future use and
- How do they express the content of these references (for instance with subject labels or summaries)?

Another conclusion from the literature is that information workers have a lot of problems with the fragmentation of information, over different collections as well as over different devices. It could be that students do not experience those difficulties to the same extent because they often work on fixed projects within a restricted period. However, over time they will probably be confronted with the same problems. Information literacy education should anticipate these problems by stimulating an

accurate attitude in document management. The same is true for the promotion of an attitude to anticipate the value of information that seems not to be valuable at that moment but that could be useful in the future. In that sense the development of PIM skills contributes to the education of life long learners.

The third conclusion from the literature is that information workers manage their information items in different ways. There is no one best approach. The preferred method(s) for organising someone's personal collections depend on personal style, the type of information that should be managed and the types of devices that are used for retrieval. The literature indicates that for computer files and emails project folders are a good solution because these items are often strongly related to the task that is being processed. But for more general documentation (not so strongly related to the task that is being executed at that moment) a flat approach might be preferred. However, when items should be retrieved with a mobile device it seems that "search" is the "smartest" way. This observation leads to my conclusion that also for information literacy curricula in Higher Education, it should be an aim to let students discover which approach(es) suits them best in different situations.

Acknowledgements

My acknowledgments go to two colleagues. The first one is Wouter Gerritsma from Wageningen UR who drew my attention to the topic PIM and William Jones' chapter on PIM in the ARIST. Peter Becker is acknowledged because he was so very kind to give his comments on the first version of this paper. They both made clear that also in Personal Information Management collaboration between people is a key factor for success.

References

ACRL. (2000). *Information literacy competency standards for higher education.* Chicago: American Library Association. Retrieved Nov. 25, 2010 from http://www.ala.org/ala/mgrps/divs/acrl/standards/standards.pdf.

Barreau, D. (2008). The persistence of behavior and form in the organization of personal information. *Journal of the American Society for Information Science and Technology, 59* (2), 307-317.

Bergman, O., Beyth-Marom, R., & Nachmias, R. (2003). The user-subjective approach to personal information management systems. *Journal of the American Society for Information Science and Technology, 54* (9), 872-878.

Bergman, O., Beyth-Marom, R., & Nachmias, R. (2008a). The user-subjective approach to personal information management systems design: evidence and implementations. *Journal of the American Society for Information Science and Technology, 59* (2), 235-246.

Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N., & Whittaker, S. (2008b). Improved search engines and navigation preference in personal information management. *ACM Transactions on Information Systems, 26* (4, 20), 1-24.

Boardman, R. (2004). *Improving tool support for personal information management*. London: Imperial College.

Boardman, R. & Sasse, M. (2004). Stuff goes into the computer and doesn't come out: a cross-tool study of personal information management. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Retrieved November 27, 2010 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.3372&rep=rep1&type=pdf.

Bruce, H. (2005). Personal, anticipated information need. *Information Research, 10* (3). Retrieved December 14, 2010 from http://informationr.net/ir/10-3/paper232.html.

Chau, D., Myers, B., & Faulring, A. (2008). What to do when search fails: finding information by association. *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems.*

Chernov, S. (2008). Task detection for activity-based desktop search. *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.*

Coughlan, T. & Johnson, P. (2009). Designing personal information management systems for creative practitioners. *Journal of Digital Information, 10* (5). Retrieved December 1, 2010 from http://journals.tdl.org/jodi.

Cushing, A. (2010). Highlighting the archives perspective in the personal digital archiving discussion. *Library Hi Tech, 28* (2), 301-312.

Dearman, D. & Pierce, J. (2008). It's on my other computer!: computing with multiple devices. *Proceedings of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems.*

Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., & Robbins, D. (2003). Stuff I've seen: a system for personal information retrieval and re-use. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.*

Elsweiler, D., Baillie, M., & Ruthven, I. (2008). Exploring memory in email refinding. *ACM Transactions on Information Systems, 26* (4, 21), 1-36.

Elsweiler, D., Baillie, M., & Ruthven, I. (2009). On understanding the relationship between recollection and refinding. *Journal of Digital Information, 10* (5). Retrieved December 1, 2010 from http://journals.tdl.org/jodi.

Elsweiler, D., Ruthven, I., & Jones, C. (2007). Towards memory supporting personal information management tools. *Journal of the American Society for Information Science and Technology, 58* (7), 924-946.

Glória, H., Guerreiro, T., & Gonçalves, D. (2010). SmarterPhone: supporting meetings in a mobile context. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries.*

Gonçalves, D. & Jorge, J. (2008). In search of personal information: narrative-based interfaces. *Proceedings of the 13th International Conference on Intelligent User Interfaces.*

Indratmo, I. & Vassileva, J. (2008). A review of organizational structures of personal information management. *Journal of Digital Information, 9* (26).

Jones, W. (2004). Finders, keepers? The present and future perfect in support of personal information management. *First Monday, 9* (3). Retrieved November 28, 2010 from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1123/1043.

Jones, W. (2007). Personal Information Management. In: Blaise Cronin (Ed.), *Annual Review of Information Science and Technology* (pp. 453-504). Medford NJ: Information Today.

McGeachin, R. (2004). The impact of electronic bibliographic databases and electronic journal articles on the scholar's information-seeking behavior and personal collection of "reprints". *Science & Technology Libraries, 25* (1/2), 127-137.

Williams, P., John, J., & Rowland, I. (2009). The personal curation of digital objects: a lifecycle approach. *Aslib Proceedings: New Information Perspectives, 61* (4), 340-363.

Yu, S. (2008). Design of a model of publication sharing and harvesting. *The Electronic Library, 26* (4), 582-593.

Content Analysis of ACRL Blog

Seda Şeker *and* **Güleda Düzyol**

Department of Information Management, Hacettepe University, 06800 Beytepe, Ankara, Turkey.
Email: {derman06, gduzyol}@hacettepe.edu.tr

Abstract: *This study analyzes ACRLog (the ACRL Blog) by its content. Using the stratified random sampling technique, 277 posts were selected from a total of 991 posts sent to ACRLog from September 2005 to November 2010. These 277 posts are grouped by subject and by author's status (academic, librarian). Differences in subjects by years and by authors are examined. In addition, word frequencies in the posts were calculated and the most used words and noun phrases determined. Differences in the use of these words and noun phrases by years and by authors are also analyzed.*

Keywords: *Blog; weblog; ACRL blog; ACRLog; content analysis.*

Introduction

A blog, also known as weblog, is a frequently updated/dynamically modified website that contains a series of date-time stamped brief entries (text, link, image, sound or video file or combination of these) usually in reverse chronological order (from newest to oldest) so that the most recent appears at the top of the blog's first page (Ewins, 2005; Schmidt, 2007; Herring et al., 2006; Tripathi & Kumar, 2010). Older entries are archived on separate pages but still accessible (Bar-Ilan, 2007). Blogs allow adding comments to the entries (Ewins, 2005; Schmidt, 2007). Long entries such as essays are not suitable for blogs (Ojala, 2004).

Blogs are worldwide accessible information channels that enable open participation in the electronic public sphere, so they are good sources of public opinion and public and private events, and they create community among people with similar interests (Ewins, 2005; Thelwall, 2007; Thompson, 2003; Bar-Ilan, 2007). They offer easy, inexpensive, flexible and interactive self-publication of content (Herring et al., 2005) that gives the author more control over the content (Schmidt, 2007). Blogs form a historical record of debates (Thelwall, 2007). There are many blog types ranging from purely topic oriented to purely personal: blogs by sectors of society, specific blog types such as warblogging, political blogs, informational blogs, photoblogs and high profile blogs (Thelwall, 2007; Bar-Ilan, 2007).

Blogging came to public awareness in the mid-1990s (Thompson, 2003). It has gained high popularity in a very short time and grown exponentially (Bar-Ilan, 2007; Herring et al., 2006). The earliest blogs appeared in 1998 (Ojala, 2004). Whereas there were only 23 known blogs at the beginning of 1999, in 2000 there were thousands of blogs (Blood, 2000). In 2003 the estimated number of blogs was half a million (Thompson, 2003), in 2005 it was about 7 million and over 70 million blogs were tracked in 2007 (Sifry, 2007). Finally, in 2008 Technorati indexed over 133 million blogs (Winn, 2008).

Parallel to the dynamic character of blogs, scholarly studies related to blogs (uses and impact of blogs, analyses of personal blogs, blogging about political topics, blogosphere, etc.) have increased (Herring et al., 2006; Schmidt, 2007). There were only 11 articles published between 1995-1999 including the term "blog". In 2003 this number was 647 (Drezner & Farrell, 2004).

Web 2.0 tools – Really Simple Syndication (RSS), Instant Messaging (IM) and blogs – are popular in libraries (Stephen, 2006; Tripathi & Kumar, 2010). Librarians were among the first to start using blogs as true communication devices (Ojala, 2004). Blogs are efficient marketing tools to inform users about library services, library events, changes, additions and other developments in library services and collections (Ojala, 2004; Bar-Ilan, 2007; Alcock, 2003; Tripathi & Kumar, 2010; Weaver, 2003). In the study of Bar-Ilan (2007), use of blogs by libraries, librarians and information specialists was analyzed and it was found that librarians use blogs to disseminate professional and general information; libraries often use blogs for announcements and to disseminate information to their users. Clyde (2003, 2004b) studied the use of blogs by libraries and discovered that libraries do not

actually use the potential of the blogs (Bar-Ilan, 2007: Clyde, 2003; Clyde, 2004a). Crawford (2005, 2006) also analysed librarian blogs. According to Tripathi and Kumar (2010), many libraries use blogs to list new databases, inform students about the downtime of servers/databases, announce hours of operation and holidays. Clyde (2004b) studied 55 library blogs using content analysis and found that they were commonly used for providing news, information and links to Internet resources for library users.

McLemee (2005) mentioned lack of a blog on academic librarianship. There are blogs by academic librarians/libraries but it is difficult to find a blog including all issues about academic and research libraries. ACRL (Association of College and Research Libraries) Strategic Plan 2020 mentions "Increasing ACRL's communication on major trends and issues in libraries..." (ACRL, 2009). ACRL Blog, called ACRLog, started in September 2005 and is a blog for academic and research librarians, sponsored by ACRL (ACRL, 2005). ACRLog's mission is "to discuss the issues of the day in the field of academic and research librarianship..." (ACRL, 2011).

This study aims to analyze ACRLog by its content. A total 991 posts were sent to ACRLog from September 2005 through November 2010, of which 277 have been selected using stratified random sampling technique. These 277 posts are grouped by subject and by author's status (academic, librarian). Differences in subjects by years and by authors and words in the posts are examined.

The following research questions are addressed by this study:

- What are the main topics of the ACRLog posts?
- Do the posts differ by subjects according to years? If this is true, which years and which subjects cause this difference?
- Do the subjects of posts differ by author's status? If this is true, which subjects are these?
- What are the most frequently used words and noun phrases in the posts? Do these mostly used words and noun phrases differ by year? If this is true, which words, noun phrases and years cause the difference?
- Do most frequently used words and noun phrases differ by author's status? If this is true, what are these words and noun phrases?
- How are the mostly used words and noun phrases scattered among the posts? Are they from a limited number of posts or from a wide range of posts?

Data and Methods

The number of posts taken into account in this study (n=277) was determined with the following formula:

$$n = 1 / ((1 / N) + (e^2 / (z^2 (p (1 - p)))) \quad (1)$$

In formula (1), the error rate (e) is 0,05, confidence level is 1,96 and significance value (p) is 0,50. The stratified random sampling technique was used with the sampling ratio values (N_i/N) as years are strata. According to the results of stratified random sampling, sample sizes for each year are shown in Table 1.

Table 1. The number of ACRLog posts by year (universe: N_i, sample: n_i)

Year	N_i	n_i	N_i/N (n_i/n)	N_i/n_i
2005	108	31	0.11	28.70
2006	294	83	0.30	28.23
2007	241	66	0.24	27.38
2008	144	42	0.15	29.16
2009	123	33	0.12	26.82
2010	82	22	0.08	26.82
Total	991	277	1.00	27.95

Table 1 also shows the number of posts by year. Notice that 2005 data only cover 4 months. 2006 has the highest number of posts and then the number of posts has decreased until 2010, which has the fewest number of posts (10 posts were sent between September-December 2010).

A sample of 277 posts were categorized under 11 subjects using the subject categorization of Järvelin and Vakkari (1993) (see Table 2). There were 16 authors registered to ACRLog. Authors were classified as academic and librarian to test the differences between their posts. Five authors were academicians and 11 authors were librarians. Using an Excel program, frequencies of all the words and noun phrases in the posts have been calculated. There were 33 words with frequencies more than 100 and 14 noun phrases with high frequencies.

Content analysis was used in this study. Content analysis methods was used earlier to analyze the structural properties and themes of blogs (Herring et al., 2006). The chi-square test was performed to test whether there is any difference in the subjects of the posts and in the use of words and noun phrases by year and by author's status. If there are differences, adjusted residual values have been examined to find the causes of these differences. Excel and SPSS programs were used for the analysis.

Findings

Findings of the study can be grouped into three categories: subjects by years, subjects by authors and words/noun phrases (by years and by authors).

Subjects by years

Table 2. Categorization of posts sent to ACRLog by subjects (2005-2010)

Subject	2005	2006	2007	2008	2009	2010	Total
The Professions	0	1	1	0	3	1	6
Library History	0	0	0	1	2	0	3
Publishing	6	2	5	3	0	0	16
Education in LIS	4	7	4	4	2	6	27
Analysis of LIS	0	2	3	1	1	0	7
LIS Service Activities	2	18	11	8	6	2	47
Information Storage and Retrieval	2	2	5	1	1	2	13
Information Seeking	3	6	0	2	2	3	16
Scientific and Professional Communication	6	9	9	7	4	2	37
Technological aspects of LIS	6	24	16	5	8	5	64
Others	2	12	12	10	4	1	41
Total	31	83	66	42	33	22	277

Table 2 shows 277 posts categorized by subjects according to the years they were posted. Most of the posts were about "technological aspects of LIS", "LIS service activities", "others" (conference announcements, information and announcements about ACRL, personal ideas and experiences, etc.) and "scientific and professional communication". By year, the posts sent in 2005 were mostly about "publishing", "scientific and professional communication" and "technological aspects of LIS". The most frequent subjects between 2006-2009 were "technological aspects of LIS", "LIS service activities" and "others". In 2010 a change was observed in the subjects of the posts; "education in LIS", "technological aspects of LIS" and "information seeking" were primary subjects for 2010.

To analyze the difference in the subjects of the posts by years, the chi-square analysis was carried out. There was a statistically significant difference in the subjects of the posts by years (p=0,017, G^2=450).[1] Adjusted residuals were examined to find which years and which subjects caused the difference exactly.

[1] As 42 cells (71.2%) have expected values bigger than 5 (20%), we used the Likelihood Ratio (G^2) test statistic instead of Pearson's Chi-Square.

Table 3. Years and subjects that cause the difference

Subject		2005	2006	2007	2008	2009	2010	Total
Profession	% Subject	0	16.7	16.7	0	**50**	16.7	100
	% Year	0	1.2	1.5	0	**9.1**	4.5	2.2
	Adjusted Residual	-0.6	-0.7	-0.4	-1	**2.9**	0.8	
Publishing	% Subject	**37.57**	12.5	31.3	18.8	0	0	100
	% Year	**19.4**	2.4	7.6	7.1	0	0	9.7
	Adjusted Residual	**3.4**	-1.6	0.7	0.4	-1.5	-1.2	
Education in LIS	% Subject	14.8	25.9	14.8	14.8	7.4	**22.2**	100
	% Year	12.9	8.4	6.1	9.5	6.1	**27.3**	9.7
	Adjusted Residual	0.6	-0.5	-1.2	-0.1	-0.8	**2.9**	
Information Seeking	% Subject	18.8	37.5	**0**	12.5	12.5	18.8	100
	% Year	9.7	7.2	**0**	4.8	6.1	13.6	5.8
	Adjusted Residual	1	0.7	**-2.3**	-0.3	0.1	1.6	
Total	% Subject	11.2	30	23.8	15.2	11.9	7.9	100
	% Year	100	100	100	100	100	100	100

Table 3 shows the cells that represent the years and subjects that show a difference in bold.[2] In 2005 "publishing", in 2007 "information seeking", in 2009 "profession" and in 2010 "education in LIS" show the differences. Posts in 2005 were primarily about "publishing" (38%). There were no posts about "information seeking" in 2007. More than half of the posts about "profession" were sent in 2009. The posts in 2010 were mostly about "profession in LIS" (27%).

Subjects by author's status

Table 4. Number of the posts by subject according to author's status

Subject	Academic	Librarian	Total
The Professions	2	4	6
Library History	0	3	3
Publishing	5	11	16
Education in LIS	15	12	27
Analysis of LIS	4	3	7
LIS Service Activities	27	20	47
Information Storage and Retrieval	12	1	13
Information Seeking	12	4	16
Scientific and Professional Communication	19	18	37
Technological aspects of LIS	40	24	64
Others	28	13	41
Total	165	112	277

Table 4 shows that 165 of the 277 posts were sent by academics and 112 posts by librarians. Both of the author groups mostly write about "technological aspects of LIS".

We also tested whether there is any difference between subjects of the posts according to the author's status and found that there is a statistically significant difference ($p=0.007$, $G^2=24.069$). Table 5 shows which subjects show the difference between academicians and librarians.

The subjects that caused the difference were "publishing" and "information storage and retrieval". Most of the posts about "publishing" were sent by librarians (69%) and almost all of the posts about "information storage and retrieval" were sent by librarians (92%).

[2] Adjusted residual values that have a z (0.95) value more than 1.96 show difference.

Table 5. Differences in posts by subjects according to the author's status

Subject		Academic	Librarian	Total
Publishing	% Subject	**31.3**	**68.8**	100
	% Year	3	9.7	5,8
	Adjusted residual	**-2.3**	**2.3**	
Information Storage and Retrieval	% Subject	**92.3**	**7.7**	100
	% Year	7.3	0.9	23.1
	Adjusted residual	**2.5**	**2.5**	
Total	% Subject	59.2	40.8	100
	% Year	100	100	100

Words and noun phrases

Table 6 shows 33 words that have frequencies greater than 100. ACRLog posts show a high frequency of the words "library" (1300) and "librarian" (712). Table 7 shows the most frequently used noun phrases in ACRLog posts, "academic librarian" (382) and "academic library" (189). Despite the rapid growth in information technologies, variation of information sources and increase in the personal information need and the increase of electronic information sources and electronic services, "information literacy" was used more than "user education".

Table 6. Words having frequencies greater than 100

Word	N	Word	N	Word	N
library	1300	blog	211	education	160
librarian	712	conference	209	college	157
student	674	article	197	access	156
academic	663	user	196	web	151
information	405	google	189	learning	136
research	328	program	186	community	135
faculty	310	post	182	higher	132
ACRL	288	resource	178	literacy	129
book	248	report	171	scholarly	120
member	223	technology	167	dijital	105
university	211	service	162	ACRLog	104

Table 7. Most used noun phrases in ACRLog

Noun phrase	N	Noun phrase	N
academic librarian	380	web 2.0	13
academic library	189	university librarian	13
higher education	75	university library	13
information literacy	70	reference services	12
search engine	51	academic freedom	12
open access	28	user education	9
web site	27	interlibrary loan	9

Although no statistically significant difference was detected in terms of number of posts between academics (165) and librarians (112), most frequently used words and noun phrases were used by academics (77% and 78%, respectively).

Words and noun phrases by year

Most frequently used words and noun phrases by year are shown in Table 8 and Table 9, respectively.

Table 8. Most frequently used words by year

Word	2005		2006		2007		2008		2009		2010		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
library	113	9	342	26	303	24	235	18	162	12	145	11	1300	100
librarian	30	4	198	28	145	20	143	20	99	14	97	14	712	100
student	46	7	161	24	134	20	108	16	108	16	117	17	674	100
academic	39	6	217	33	143	22	119	18	76	11	69	10	663	100
information	40	10	118	29	86	21	69	17	60	15	32	8	405	100
research	22	7	65	20	58	18	71	22	59	18	52	16	328	100
faculty	16	5	85	27	40	13	53	17	73	24	43	14	310	100
ACRL	18	6	122	42	81	28	21	7	32	11	14	5	288	100
book	31	13	34	14	66	27	80	32	26	10	11	4	248	100
member	32	14	100	45	35	16	24	11	26	12	6	3	223	100
university	22	10	53	25	35	17	52	25	31	15	18	9	211	100
blog	20	9	43	20	64	30	31	15	44	21	9	4	211	100
conference	29	14	71	34	39	19	26	12	11	5	33	16	209	100
article	17	9	58	29	33	17	26	13	25	13	39	20	197	100
user	16	8	85	43	27	14	16	8	12	6	40	20	196	100
google	26	14	75	40	35	19	22	12	15	8	16	9	189	100
program	22	12	78	42	15	8	27	15	30	16	14	8	186	100
post	10	5	40	22	25	14	60	33	35	19	13	7	182	100
resource	17	10	51	29	33	18	30	17	21	12	26	15	178	100
report	13	8	84	49	29	17	17	10	11	6	17	10	171	100
technology	5	3	56	34	42	25	20	12	39	23	5	3	167	100
service	9	6	49	30	28	17	20	12	21	13	35	22	162	100
education	27	17	58	36	34	21	14	9	19	12	8	5	160	100
college	9	6	46	29	27	17	29	18	31	20	15	10	157	100
access	3	2	30	19	30	19	18	12	38	24	37	24	156	100
web	21	14	45	30	45	30	9	6	20	13	11	7	151	100
learning	20	15	33	24	36	26	16	12	19	14	12	9	136	100
community	9	7	33	24	24	18	29	21	30	22	11	8	135	100
higher	10	8	41	31	32	24	18	14	20	15	11	8	132	100
literacy	29	22	45	35	7	5	21	16	22	17	5	4	129	100
scholarly	1	1	17	14	21	18	25	21	39	33	17	14	120	100
dijital	11	10	31	30	27	26	19	18	7	7	10	10	105	100
ACRLog	3	3	29	28	36	35	23	22	10	10	3	3	104	100

"Library", "librarian", "student", "academic", "ACRL" "education", "web", "google" and "technology" were the most frequently used words in 2006 and 2007 while "university", "research" and "blog" were more prevalent in 2006 and 2008 (Table 8). "Book" was used most often in 2007 and 2008, "literacy" in 2006 and 2009, "user" in 2007 and 2010 and "access" in 2009 and 2010. "Academic librarian" was the most frequently used noun phrase for each year (Table 9). "Academic library" was the most frequently used noun phrase in 2006 (29%) and 2007 (27%); "information literacy" in 2006 (30%) and 2008 (21%), "search engine" in 2006 (54%), "open access" in 2009 (28%) and 2010 (35%), and "web site" in 2006 (40%). The use of the noun phrase "open access" in the last two years shows that this subject gained importance. The decrease in the use of the noun phrase "information literacy" in recent years is an interesting finding because there have been lots of studies on this subject. One of the other interesting findings is the use of "Web 2.0" in 2005 and 2006 and no use in recent years since Web 2.0 has increasing importance for university and research libraries (Tripathi & Kumar, 2010).

Table 9. Most used noun phrases by years

Noun phrases	2005		2006		2007		2008		2009		2010		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
academic librarian	20	5	118	31	76	20	83	22	42	11	41	11	469	100
academic library	11	6	54	29	50	27	28	15	22	12	24	13	278	100
higher education	7	9	26	35	15	20	10	13	13	17	4	5	169	100
information literacy	11	16	21	30	4	6	15	21	13	18	6	9	161	100
search engine	5	10	28	55	10	20	4	8	-	-	4	8	144	100
open access	1	4	2	7	5	18	2	7	8	29	10	36	93	100
web site	6	22	11	41	6	22	3	11	1	4	-	-	127	100
web 2.0	4	31	2	15	5	39	-	-	2	15	-	-	113	100
university librarian	1	15	7	31	-	8	5	46	-	-	-	-	113	100
university library	2	8	4	54	1	-	6	39	-	-	-	-	114	100
reference services	2	17	1	8	4	33	5	42	-	-	-	-	112	100
academic freedom	2	17	-	-	6	50	4	33	-	-	-	-	112	100
user education	4	44	5	57	-	-	-	-	-	-	-	-	110	100
interlibrary loan	3	33	1	11	2	22	-	-	-	-	3	33	75	100

The most frequently used words and noun phrases by years exhibit statistically significant differences ($p=0.000$, $X^2=928.279$; $p=0.000$, $G^2=206.274$).

Words and noun phrases by authors' status

The most frequently used words and noun phrases by authors' status are shown in Table 10 and 11.

Table 10. Most used words according to author's status (academic, librarian)

Word	Academic	Librarian	Total	Word	Academic	Librarian	Total
library	995 (1)	305 (1)	1300	post	132 (17)	50 (10)	182
librarian	557 (3)	155 (3)	712	resource	127 (19)	28 (19)	178
student	475 (4)	199 (2)	674	report	124 (20)	47 (12)	171
academic	573 (2)	90 (6)	663	technology	143 (15)	24 (21)	167
information	279 (6)	126 (5)	405	service	113 (23)	49 (11)	162
research	287 (5)	41 (14)	328	education	139 (16)	21 (23)	160
faculty	261 (8)	49 (11)	310	college	116 (21)	41 (15)	157
ACRL	270 (7)	18 (24)	288	access	98 (27)	58 (9)	156
book	97 (28)	151 (4)	248	web	130 (18)	21 (23)	151
member	208 (9)	15 (25)	223	learning	108 (25)	28 (19)	136
university	151 (13)	60 (8)	211	community	110 (24)	25 (20)	135
blog	166 (11)	45 (13)	211	higher	114 (22)	18 (24)	132
conference	188 (10)	21 (23)	209	literacy	106 (26)	23 (22)	129
article	132 (17)	66 (7)	197	scholarly	89 (29)	31 (18)	120
user	156 (12)	40 (16)	196	digital	60 (31)	45 (13)	105
google	139 (16)	50 (10)	189	ACRLog	83 (30)	21 (23)	104
program	147 (14)	39 (17)	186				

The most frequently used word by both author groups was "library". This was followed by "academic", "librarian", "student" and "research" by academics, and "student", "librarian", "book" and "information" by librarians. It is interesting to note that the word "digital" is used least by academics even though with the developing technology most information sources are in digital format nowadays. On the other hand, the word "book" was used more frequently by librarians than by academics.

The noun phrases "academic library" and "academic librarian" were used most often by both academics and librarians. Librarians used "information literacy", "search engine" and "open access" a more frequently while academics used "higher education" more often. "Information literacy" was preferred more frequently than "user education" by both academics and librarians.

Table 11. Most frequently used noun phrases by author's status (academic, librarian)

Noun phrases	Academic	Librarian	Total
academic librarian	320	60	380
academic library	153	36	189
higher education	68	7	75
information literacy	45	25	70
search engine	26	25	51
open access	11	17	28
web site	24	3	27
web 2.0	9	4	13
university librarian	12	1	13
university library	11	2	13
reference services	5	7	12
academic freedom	8	4	12
user education	9	0	9
interlibrary loan	3	6	9

The most frequently used words and noun phrases differ statistically significantly by author's status. ($p=0.000$, $X^2=506.095$; $p=0.000$, $G^2=87.483$) (see Table 12 and Table 13). All words except "book" "book" were used more often by academics, indicating that academics use more words than librarians do in their posts. The noun phrases "open access", "reference service", "interlibrary loan" were used more often by librarians, and "academic librarian", higher education", "information literacy" by academics.

Table 12. Words that show a difference between authors

Word	Academic		Librarian		Total	
	N	%	N	%	N	%
student	475	71	199	29	674	100
academic	573	86	90	14	663	100
information	279	69	126	31	405	100
research	287	88	41	12	328	100
faculty	261	84	49	16	310	100
ACRL	270	94	18	6	288	100
book	97	39	151	61	248	100
member	208	93	15	7	223	100
university	151	72	60	28	211	100
conference	188	90	21	10	209	100
article	132	67	66	33	197	100
service	113	70	49	30	162	100
education	139	87	21	13	160	100
access	98	63	58	37	156	100
web	130	86	21	14	151	100
higher	114	86	18	14	132	100
dijital	60	57	45	43	105	100

Table 13. Noun phrases that show difference between authors

Noun phrase	Academic		Librarian		Total	
	N	%	N	%	N	%
academic librarian	320	84	60	16	380	100
higher education	68	91	7	9	75	100
information literacy	45	64	25	36	70	100
search engine	26	51	25	49	51	100
open access	11	39	17	61	28	100
reference services	5	42	7	58	12	100
interlibrary loan	3	33	6	67	9	100

Appendix shows scattering of most frequently used words to the posts. The word "library" and "librarian" were used by most of the authors (213 and 202 post have "library" and "librarian" in them, respectively). The word "academic" came from 185 different posts, and "university" from 100 different posts. The word "conference" is used 209 times but this frequency comes from 70 posts. Similarly the word "ACRL" is used 288 times in 95 different posts. Despite frequent use of the word "google", relatively fewer number of posts (57) included "google". Appendix shows that the use of words in ACRL blog posts is widely scattered.

Conclusion

We performed a content analysis with 277 posts sent to ACRLog. Most of the posts were about "technological aspects of LIS", "LIS service activities", "others" (conference announcements, information and announcements about ACRL, personal ideas and experiences, etc.) and "scientific and professional communication". There is a statistically significant difference in the subjects of the posts by year. Some 33 words have frequencies of more than 100. ACRLog posts mostly contain the words "library" (1300) and "librarian", and the noun phrases "academic librarian" and "academic library". Scattering of words to the posts is wide. Some of the words like "ACRL", "conference", "google", etc. come from certain blog posts. The subjects of blog posts and most frequently used words/noun phrases in the blog posts provide consistent results.

To increase the quality of ACRL blog as a professional communication, the interest in the ACRL blog should be increased and current issues such as Web 2.0, information literacy and digital rights should be discussed more often to enrich the content

References

ACRL. (2005). Why an ACRL Blog. *College & Research Libraries News, 66(10), 725*. Retrieved January 23, 2011 from http://crlnews.highwire.org/content/66/10/725.full.pdf+html.

ACRL. (2009). *Charting our future: ACRL Strategic Plan 2020*. Retrieved January 23, 2011 from http://www.ala.org/ala/mgrps/divs/acrl/about/whatisacrl/strategicplan/index.cfm.

ACRL. (2011). *Why an ACRL Blog*. Retrieved January 23, 2011 from http://acrlog.org/why-an-acrl-blog/.

Alcock, M. (2003). Blogs – what are they and how do we use them? *Quill*, 103(8). Retrieved January 21, 2011 from http://alia.org.au/groups/quill/issues/2003.8/blogs.html.

Bar-Ilan, J. (2007). The use of Weblogs (blogs) by librarians and libraries to disseminate information. *Information Research*, 12(4). Retrieved January 22, 2011 from http://InformationR.net/ir/12-4/paper323.html.

Blood, R. (2000) Weblogs: a history and perspective, *Rebecca's Pocket*. Retrieved January 19, 2011 from http://www.rebeccablood.net/essays/weblog_history.html.

Clyde, L.A. (2003). Weblogs in the library and information science environment. *Paper presented at the Online 2003 Conference*, London.

Clyde, L.A. (2004a). *Weblogs and libraries*. Oxford: Chandos Publishing.

Clyde, L.A. (2004b). Library weblogs. *Library Management*, 25(4-5), 183-189. Retrieved January 21, 2011 from http://mysite.pratt.edu/~dwalczyk/clyde.pdf.

Crawford, W. (2005). Investigating the biblioblogosphere. *Cites & Insights*, 5(10). Retrieved January 21, 2011 from http://citesandinsights.info/v5i10b.htm.

Crawford, W. (2006). Looking at liblogs: The great middle. *Cites & Insights*, 6(10). Retrieved January 21, 2011 from http://citesandinsights.info/v6i10a.htm.

Drezner, D. & Farrell, H. (2004, September). The power and politics of blogs. *Proceedings of the annual meeting of the American Political Science Association*. Chicago, IL: American Political Science Association. Retrieved January 20, 2011 from http://www.danieldrezner.com/research/blogpaperfinal.pdf

Ewins, R. (2005). Who are you? Weblogs and academic identity. *E-learning and Digital Media*, 2(4), 368-377. Retrieved January 18, 2011 from http://dx.doi.org/10.2304/elea.2005.2.4.368.

Herring, S.C., Scheidt, L.A., Kouper, I., & Wright, E. (2006). Longitudinal content analysis of weblogs: 2003–2004. In M.Tremayne (Ed.), *Blogging, Citizenship, and the Future of Media* (pp. 3–20). London: Routledge.

Herring, S.C., Scheidt, L.A., Bonus, S., & Wright, E. (2005). Weblogs as a bridging genre. *Information,*

Technology & People, 18(2), 142–171.

Järvelin, K. & Vakkari, P. (1993). The evolution of library and information science 1965-1985: A content analysis of journal articles. *Information Processing and Management*, 29(1), 129-144.

McLemee, S. (2005). Silence in the stacks, *Inside Higher Ed,* Retrieved January 23, 2011 from www.insidehighered.com/views/2005/06/09 /mclemee.

Ojala, M. (2004). Weaving Weblogs into knowledge sharing and dissemination. *NORD I&D, Knowledge and Change*, 212-220.

Schmidt, J. (2007). Blogging practices: An analytical framework. *Journal of Computer-Mediated Communication*, 12(4). Retrieved January 21, 2011 from http://jcmc.indiana.edu/vol12/issue4/schmidt.html.

Sifry, D. (2007, April). The state of the live web, April 2007. *Technorati Weblog*. Retrieved January 20, 2011 from http://www.sifry.com/alerts/archives/000493.html

Stephens, M. (2006). Exploring Web 2.0 and libraries. *Library Technology Reports*, 42(4), 8-14.

Thelwall, M. (2007). Blog searching: the first general-purpose source of retrospective public opinion in the social sciences?, *Online Information Review*, 31(3), 277-289.

Thompson, G. (2003). Weblogs, warblogs, the public sphere, and bubbles, *Transformations*, 7, Retrieved January 21, 2011 from http://www.transformationsjournal.org/journal/issue_07/article_02.shtm.

Tripathi, M. & Kumar, S. (2010). Use of Web 2.0 tools in academic libraries: A reconnaissance of the international landscape. *The International Information & Library Review*, 42 (3), 197-207. Retrieved January 20, 2011 from ScienceDirect database.

Weaver, B. (2003, September). Weaver's web. *InCite*. Retrieved January 21, 2010, from http://archive.alia.org.au/incite/2002/09/weaver.html.

Winn, P. (2008). *State of the blogosphere: Introduction 2008*. Retrieved January 21, 2011, from http://technorati.com/blogging/article/state-of-the-blogosphereintroduction/.

Appendix: Most frequently used words' scattering to the posts (N_2) and frequencies of most used words (N_1)

Word	2005		2006		2007		2008		2009		2010		Total	
	N_1	N_2	N_1	N_2	N_1	N_2	N_1	N_2	N_1	N_2	N_1	N_2	N_1	N_2
library	113	24	342	60	303	47	235	30	162	30	145	22	1300	213
librarian	30	23	198	56	145	44	143	32	99	25	97	22	712	202
student	46	13	161	37	134	31	108	22	108	18	117	16	674	137
academic	39	19	217	57	143	39	119	30	76	23	69	17	663	185
information	40	18	118	48	86	32	69	21	60	15	32	9	405	143
research	22	13	65	29	58	24	71	20	59	21	52	14	328	121
faculty	16	12	85	30	40	18	53	21	73	15	43	14	310	110
ACRL	18	9	122	31	81	23	21	15	32	11	14	6	288	95
book	31	9	34	22	66	23	80	23	26	14	11	9	248	100
member	32	5	100	29	35	14	24	12	26	11	6	8	223	79
university	22	12	53	29	35	20	52	17	31	12	18	10	211	100
blog	20	9	43	24	64	19	31	19	44	14	9	10	211	95
conference	29	8	71	18	39	16	26	16	11	8	33	4	209	70
article	17	8	58	29	33	17	26	16	25	12	39	13	197	95
user	16	7	85	19	27	17	16	8	12	5	40	5	196	61
google	26	9	75	12	35	9	22	10	15	7	16	10	189	57
program	22	11	78	26	15	12	27	13	30	16	14	6	186	84
post	10	11	40	31	25	21	60	24	35	19	13	8	182	114
resource	17	6	51	27	33	17	30	14	21	10	26	10	178	84
report	13	7	84	25	29	11	17	8	11	15	17	9	171	75
technology	5	6	56	19	42	17	20	10	39	11	5	2	167	65
service	9	7	49	26	28	16	20	18	21	11	35	10	162	88
education	27	18	58	32	34	17	14	11	19	8	8	6	160	92
college	9	3	46	25	27	12	29	15	31	13	15	9	157	77
access	3	4	30	18	30	19	18	12	38	11	37	13	156	77
web	21	8	45	18	45	18	9	14	20	6	11	8	151	72
learning	20	10	33	22	36	17	16	11	19	7	7	7	136	74
community	9	4	33	20	24	14	29	16	30	11	11	16	135	81
higher	10	9	41	22	32	15	18	11	20	9	11	8	132	74
literacy	29	8	45	19	7	3	21	9	22	8	5	5	129	52
scholarly	1	1	17	8	21	12	25	9	39	13	17	4	120	47
dijital	11	8	31	13	27	11	19	8	7	4	10	5	105	49
ACRLog	3	2	29	21	36	14	23	9	10	7	3	3	104	56

An Introduction to M-Learning in Iranian Universities

Mehdi Afzali *and* **Mahdi Bazargani**

Department of IT Engineering, Zanjan Branch,Islamic Azad University, Zanjan, Iran,
Email: afzali@hacettepe.edu.tr, mbzir@yahoo.com

Abstract*: The first experience of remote education in Iran was started in 1970 by the Open University of Iran. In 1991, an e-learning site in Tehran University provided nine courses for students for the first time. Currently, the majority of Iran's universities offer these courses. To benefit more from modern facilities and technology, recently some universities have attempted to use Mobile Learning (M-Learning) based education. One of the most important challenges which these universities encounter is the lack of digital information resources and the difficulty of preparing digital content in the Persian language. In this research, services related to M-Learning are considered along with technical impediments in the development of these services in Iran.*
Keywords*: M–learning; Persian content; virtual universities in Iran.*

Introduction

Increasing the use of ICT in education has led to changes in methods, systems and resources. Migration of information resources from print mode to electronic mode has provided specific advantages for researchers and has changed the exploitation of information resources. In recent decades the development of education in a virtual environment and use of new tools have caused new methods of training to emerge.

The cellular phone is a telecommunication tool which is available in most parts of the world, both developed and developing countries, and is affordable for everyone. In this situation, it is possible to think about other applications of the cellular phone. Sending text is possible with almost any type of cellular phone and operator, and in more advanced models, sending audio and video files is also feasible. Even text files can easily include educational topics. Accordingly, utilization of mobile learning requires no expensive equipment.

Today, by creating and developing new information technologies, education for everyone, everywhere and at any time would seem to be obvious. On the other hand, attention to education in the modern world has persuaded owners of industries and technologies to innovate and provide education along with common telephone services. The complexity of the technology offered with cellular phones on one hand, and focusing on the educational needs of the applicants on the other hand, has become one of the most basic concerns of mobile technology custodians.

M-Learning implies a learning environment in which instructors and students can access the learning system with portable devices over a wireless network. The educational environment is becoming better than ever in ubiquity, instant connectivity, personalization, and a self-directed learning community with mobile technology (Kwon and Lee, 2010, p. 1885).

Here, the meaning of mobile learning is training by means of digital mobile tools which are portable and useable anywhere and anytime. These tools include:

- Mobile Telephones;
- PDAs (Personal Digital Assistants); and
- Personal Digital Media Players.

Although laptop computers can also be considered part of this group, our discussion focuses on the small screens with no standard keypad.

The definition of mobile learning has evolved with the advent of new technology. While mobile learning could, in its broadest sense, be said to cover books, CD-ROMs, radios, and laptops, most researchers in the field of educational technology consider mobile learning, or M-Learning, to be a

subset of e-learning. What differentiates mobile from electronic is the nature of the technology. M-learning terms, however, exhibit a shift away from a media immersion experience, or a structured, benchmarked set of activities, to words that express what are perceived as the primary attributes of a very different experience: "spontaneous, intimate, situated, connected, informal, lightweight, personal" (Laouris & Eteokleous, 2005, 3). S. Geddes (2004, p. 1) has provided a succinct definition: "mLearning is the acquisition of any knowledge and skill through using mobile technology, anywhere, anytime, that results in an alteration in behaviour".

For the purposes of this paper, mobile learning denotes instructional content or activities that are delivered on handheld (or mobile) devices, that accommodate limited multimedia delivery, primarily in the form of audio, images, animations (video), and text. Popular mobile devices include those capable of playing files that are often downloaded from the Internet on a computer and then uploaded onto the device which is then taken with the individual, who can play the files while not directly connected to a computer. Audio and video files can be played on handheld computers, audio file players such as iPod, and handheld devices that play video files.

M-Learning Technologies

E-Learning technologies were presented back in 1993 when bounded computer/email-based interactions were available (Electronic Learning, 2011). E-Learning, flexible learning, and distance learning definitions often go hand-in-hand and the goal of such learning methods is to give highest location independence and flexibility to get learning materials through online and media-based (e.g., CD and DVD) technologies.

The emerging web technologies have had profound effects on the quality of e-learning methods. Web 1.0 is a term referring to the early concepts of content delivery using basic web design technologies, including; HTTP (Hypertext Transfer Protocol) and HTML (Hyper Text Markup Language). Web 2.0 is a more recent term referring to the next generation of web-related applications and content delivery and presentations that facilitates secured multimedia-enabled information sharing. Web 2.0 includes the following concepts: Social Work 2.0, online interactions (e.g., blogs), XML (Extensible Markup Language), RSS (Really Simple Syndication) feeds, social networking protocols (e.g., FOAF "Friend of a Friend" and XFN "XHTML Friends Network"), and Web APIs (Application Programming Interfaces, such as REST "Representational State Transfer" and SOAP "Simple Object Access Protocol").

Mobile computing and communication services are spreading rapidly. Research suggests that in 2009 there were nearly 250 million wireless data-capable devices in use in the United States (Paczkowski, 2009). Adoption rates for mobile technology dwarf those for non-mobile technologies; for example, there are eight times more iPhone/iPod Touch users two years after their launch than there were AOL users two years after its launch (Meeker et al, 2009). At the end of 2009, there were 4.6 billion mobile cellular subscriptions worldwide, representing two-thirds of the world population (International Telecommunication Union, 2010). Mobile industry analysts suggest that worldwide mobile data traffic will double every year through 2013, increasing 66 times between 2008 and 2013 (Cisco System, Inc, 2010).

M-Learning is on the same track as e-learning, however it features distinct device/application utilization, which is based on mobile or other types of portable devices in the e-learning process. M-Learning is completely location independent, which means that the M-Learning session can take place virtually in any location (e.g., taxi, restaurant, street, etc.) (Mobile Learning, 2011).

As cell-phone data-plans continue to decrease in price and increase in versatility and include more affordable contract terms (i.e., unlimited data usage), many people are considering M-Learning more attractive than e-learning, due to the fact that e-learning scenarios require the learners to be in a fixed location or with a limited mobility throughout the data learning delivery, mostly through the Internet. There are however a few challenges associated with M-Learning, including: limited battery life on mobile devices, variable wireless coverage and having limited or no coverage in some rural areas, and limited key and screen sizes (Mobile Learning, 2011).

In 1988, in the U.S.A. a powerful software program that represented a revolutionary change in the electronic system, entitled "Master Digital (Digital-Professor)", that offered one of the primary uses of phone computers in education, was used for the first time.

The M-Learning project was funded in 2001 by the European Commission's Information Society Technologies (IST) initiative with matched funding from the project partners and, in the UK, the Learning and Skills Council. There are five project partners: two university-based research units (Ultralab at Anglia Polytechnic University in the UK and Centro di Ricerca in Matematica Pura ed Applicata (CRMPA) at the University of Salerno in Italy), two commercial companies (Cambridge Training and Development Limited (CTAD) in the UK and Lecando in Sweden) and the Learning and Skills Development Agency (LSDA) in the UK (Attewell, 2004, p. 2).

Learner access to M-Learning project systems and materials was via a microportal (mPortal), which consists of a series of mini web pages with navigation pointing to:

- learning materials;
- mini web page builder tools;
- a collaborative activities tool (the MediaBoard);
- peer-to-peer communication services (messages, chat, discussion and blogs);
- the learning management system;
- simple help guides for the system; and
- links to places on the Web that may be helpful or interesting for the target audience (e.g. alcohol, drugs and sexual health advice services, job hunting and online learning services and dictionaries) (Attewell, 2004, p. 7).

Current Smartphone platforms, such as in Blackberry Storm and Apple iPhone, have remedied the screen and key sizes to some extent and the new technologies in the wireless domain fronts, including usage of localized cell coverage (i.e., femtocell and directing cellular traffic through the Wi-Fi Access Point (AP) using UMA (Unlicensed Mobile Access) technology, have also eased the wireless coverage issues slightly. The battery life issue is still the major challenge, though several power-save operation modes have been introduced and implemented in various cellular and Wi-Fi protocols, including: UAPSD (Unscheduled Automatic Power Save Delivery) in IEEE 802.11e and SMPS (Spatial Multiplexing Power Save) and PSMP (Power Save Multi Poll) in IEEE 802.11n (Adibi, 2010, p. 379).

E- Learning in Iran

The first experience of remote education in Iran was started in 1970 by the Open University of Iran (Peyam Noor University, 2011). In 1991, for the first time an e-learning site in Tehran University provided nine courses for students.

In 2006 the state of e-learning and application of information technology in Iran was reviewed. According to that research 15 universities in seven educational groups accepted students with a total of 1000 registered students that year. According to the research results three virtual universities received establishment licenses in the field of e-learning. In that year there was no M-Learning. (Montazer, 2006, pp. 126-127).

In Iran, use of the Web for learning with PCs is more than ten years old. During this period, the spread of e-learning was not so great. Only 13 universities and five Institutes admit students and train them for B.A./B.S. and M.A./M.S. degrees in different fields making use of e-learning. Considering the large number (more than 230) of public and private universities in Iran, the number offering e-learning is very small. The Ministry of Education also organized e-learning in 576 training centers. The eagerness of the students and their need for e-learning force universities to develop their educational services.

Mobile phone penetration has increased in Iran in recent years, as well as in the world. It is natural that the number of operations and businesses in the electronic environment has also risen. Trust in the electronic environment has become vital, so information and communication security is expected more than before.

The three mobile operators in Iran are as follow:

- Irancell Co. (www.irancell.ir);
- Mobile Communication Co. (www.mci.ir);
- Taliya Co. (www.taliya.ir).

Considering the increasing number of cellular phone users and the possibility of usage of GPRS by these systems, features such as sending Mobile Information have led to the adoption of this facility by educational institutions. The number of universities officially using M-Learning is four. Some other universities are also exploring the use of M-Learning in their educational programs.

One can divide the major problems of Iranian universities in e-learning into three groups:

- technologies related to e-learning services;
- lack of resources in the Persian context; and
- security in information exchange.

Mobile Information Services

Features like SMS and Bluetooth are among the most common information swap services offered by cellular phones with widespread use among the Iranian youth. Some news agencies have tried to give information to their audience by means of these services, for example: Iran Book News Agency (www.ibna.ir/).

Recording and playing back the audio files and the video resources has recently become common in the e-learning system in Iran. Sources which in the past were on video cassettes or CD or DVD now are in the form of digital files and are able to be displayed on PDAs, Digital Multimedia Players or cellular phones.

The outsourcing of e-books, audio books, movies, and other multimedia content presents both benefits and challenges to libraries. With school, college, and university libraries facing severely constrained budgets, only the largest and most financially stable library systems have the resources to maintain in-house a full array of technical services, including the online public access catalog (OPAC), data store, e-mail, and web servers.

Learning foreign languages, especially English, through cellular phones is also one of the issues to be considered favorably in Iran. Some of the private language institutes use the cellular phone to send educational materials, including voice, image, films and texts. As the display screen is very small in cellular phones in comparison with PCs, the amount of information that is displayed is a disadvantage. So, if the information is more than one page, the format of the display should be changed, for example, diagram, flowchart and video formats can be used. Sending audio English to Persian dictionaries is the most common use of these resources.

Portable Documents are one of the major resources to display information for trainees. Books are an important example in this field. Because the use of Word Processing and desktop publishing has increased, the growth of production in digital documents in the Persian language has been significant. At the present time, several methods are available for cellular phone users to download a document. The volume of the content and the size of display are also important, so use of multimedia resources is not much favored due to the high volume, which causes problems in sending.

The noteworthy point is that educational contents should be designed for minimum equipment in order to be feasible on most regular phones and in the areas where cellular phone operators do not support all the features.

M-Learning in Iranian Universities

A number of Iranian universities attempt to provide their students with M-Learning, as follow:

1. University of Hadith Sciences[1]

The relatively good cost of classes and also the opportunity to attend general classes free of charge are notable points of this university. Online contents and CDs are available to students. In addition to pictures and slides, the CDs also include audio courses suitable for use in mobile phones and audio devices.

[1]. http://vu.hadith.ac.ir/

Communication with teachers is done via e-mail and text chat. Physical presence in the building during examinations is obligatory. The number of information sources provided is a negative aspect in the virtual education of this university.

2. University of Shiraz[2]

Access to the education system through cell phone is a positive point of the university. Materials are available for students as CDs. Communication with teachers is via e-mail and text chat. Physical presence in the building during examinations is necessary. Lack of simultaneous web-based classes and high cost of web classes are negative points of e-learning in the university despite its nature and the quality of resources and services presented.

3. Amirkabir University of Technology (Tehran Polytechnic) [3]

The Institute of Virtual Studies including four online classes is equipped with different facilities for e-learning used for simultaneous online courses. M-Learning facilities and conditions in this university are like Shiraz University. This system is a complement for e-learning at the University.

4. Medical University of Tehran[4]

This university is the only university offering e-learning for medical sciences. In addition to the typical M-Learning of universities, they have prepared special textbooks for mobile devices. Most of these information sources have been prepared by Java programming.

5. Other Universities

Using M-Learning is under study in many public and private universities of Iran. These studies are most common among courses which do not need practical classes or workshops. Among these universities some are mentioned below: Sharif Technical University[5], Tehran University[6], Islamic Azad University[7], and Azerbaijan Virtual University.

By reviewing the services by Iranian universities via M-Learning we noticed that these services had been concentrated mostly on representation of curriculum content in the form of text and audio. Because of technological and infrastructural problems, the use of images, video and multimedia are not taking into account properly and in practice there is no application of them. Only two universities are using images for curriculum content in a limited manner.

In the field of learning assessment and exams, application of distance education systems is not common and students should be present physically during the exam period. In three universities an assessment system from a distance has been established for M.A./M.S. students.

Social science is more prominent in e-learning than other majors and favorite subjects Owing to the simplicity of preparing curriculum resources in this field. In medical science and in the arts still there is no activity. The University of Islamic Arts has begun a vast plan in the field of electronic education. Also, this university may start activities in M-Learning in the future.

---

2. http://www.shirazu.ac.ir/
3. http://www.aut.ac.ir/
4. http:// www.tums.ac.ir/
5. http://cw.sharif.ir/
6. http://ut.ac.ir
7. http://www.iau.ac.ir/

Table 1. M-Learning methods and models used in Iranian universities

		U. of Hadith Sciences	U. of Shiraz	Amirkabir U.	Medical U. of Tehran	Sharif Technical U.	Tehran U.	Islamic Azad U.	Azerbaijan Virtual U.
Methods used for content delivery	Text	x	x	x	x	x	x	x	x
	Voice	x	x	x	x	x	x	x	x
	Image				x	x			
	Film								
	Multimedia								
Methods used for e-assessment	presence	x	x	x	x	x	x	x	
	distance			x		x			x
Departments	Engineering		x	x		x			x
	Medical Science				x				
	Social Science	x	x				x	x	x

Conclusion

Cell phones play an important role in the lives of young people. They are used as a tool in voice communications and data transferring. It seems true at first glance that mobile learning is like using personal computers (only smaller in size), but our research findings indicated that in fact the technological limitations and lack of control over how and when training happens need different educational models.

There is a big gap between Iran and developed countries in technological issues. This lack of technological development has had a negative impact on the educational system. Lack of the necessary structures for telecommunication like cable broadcasting in some zones of the country has caused a more rapid growth of wireless infrastructures than in most countries in the world. Because of the low cost and wide range of cellular phone stations, more rural zones have been covered by wireless networks (Wireless LANs). Due to the vast application of cellular phones, the use of Internet has spread in many urban regions in Iran and now is available in all cities.

According to the investigations on the requirements, necessity, versatility, and projects done in Iran, the following subjects can be concluded about the problems. Schemes to solve the problems are then mentioned.

- *Lack of appropriate technological and networking infrastructures:* due to lack of proper infrastructure in rural and urban areas, the possibility of M-Learning with good quality is currently impossible and that is why virtual universities prefer not to provide these services.
- *Cultural problems in the society:* Not welcoming cellular phones as a means of sending non-audio data, and inability of most people to use this tool in education.
- *Unfamiliarity of M-Learning:* Research and investigation about M-Learning are not carried out in all universities and this system is still unknown for most of them.
- *Limited number of professors and experts:* Despite the increasing rate of growth of virtual learning and virtual universities, we still have a limited number of experts and professors who have experience in teaching students or in establishing and maintaining virtual universities. This

problem is quite palpable and has caused a tangible fall in quality of educational content, level of education and information literacy.

- *Problems related to learning and educational management systems:* Although examples of learning management systems for M-Learning have been produced already in the country, due to the lack of these systems and internal software and also their high cost, most universities and higher education institutions tend to use free foreign versions, which causes difficulty in maintenance, trouble shooting and working with these systems. Working with them is complex and time consuming, and security risks and potential security holes are inevitable.
- *Limitation of subjects and quality of electronic content in Persian language:* slow and time-consuming production of new electronic content in Farsi in the universities has led to lack of information resources. So, it leads the universities to use available resources in English.
- *Lack of support from top management:* top managers in all universities whether private or public do not look at this phenomenon (M-Learning) from a positive point of view and this has led to its slow development and weak educational system.
- *Poor or difficult educational content:* The educational material produced, despite the time-consuming and costly process, does not have an acceptable quality. It is so weak and superficial in content that the students are forced to use supplementary books, or is full of complicated phrases and too difficult due to poor translation so that students have to find alternative books or use original ones, if available.
- *Lack and non-use of national standards:* Despite existing international standards, we still do not have any electronic standard within the country to impose the use of M–Learning. Currently every university follows its own procedures and there is no centralized standard to synchronize universities for information transfer between universities.

Due to the various problems mentioned, the most important strategies to expand access and lower the limitations of M-Learning are as follow:

- Development of national standards for sending Mobile Information;
- Developing a culture of making optimal use of cell phones;
- Development of mobile learning in undergraduate levels;
- Enhancement of the software and hardware facilities;
- Preparation of required training content by professors and experts; and
- Reducing training costs.

References

Adibi, S. (2010). A remote interactive non-repudiation multimedia-based M-Learning system. *Telematics and Informatics*, 27, 377–393.

Amirkabir University of Technology. (2011). Retrieved March 20, 2011 from http://www.aut.ac.ir/

Attewell, J. (2004). *Mobile technologies and learning: A technology update and m-learning project summary*. London: Learning and Skills Development Agency. Retrieved March 20, 2011 from http://www.LSDA.org.uk

Cisco Systems, Inc., (2010). Cisco Visual Networking Index: Forecast and Methodology, 2008-2013. Retrieved March 20, 2011 from http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_ Networking_ Solutions_ White_Paper.html

Electronic Learning. (2011). *Wikipedia*, Retrieved March 20, 2011 from http://en.wikipedia.org/wiki/Electronic_learning

Geddes, S. (2004). Mobile learning in the 21st century: Benefit for learners. *Knowledge Tree e-journal*. Retrieved March 19, 2011, from https://olt.qut.edu.au/udf/OLTCONFERENCEPAPERS/gen/static/papers/Cobcroft_OLT2006_paper.pdf.

International Telecommunication Union. (2010). Measuring the Information Society 2010. Retrieved March 23, 2011 from http://www.itu.int/ITU-D/ict/publications/idi/2010/Material/MIS_2010_Summary_E.pdf

Iran Book News Agency. (2011). Retrieved March 25, 2011 from http://www.ibna.ir/

Kwon, S. & Lee, J.E. (2010). Design principles of m-learning for ESL. *Procedia – Social and Behavioral Sciences*, 2, 1884-1889.

Laouris, Y. & Eteokleous, N. (2005). We need an educationally relevant definition of mobile learning. Retrieved March 19, 2011, from http://www.mlearn.org.za/CD/papers/Laouris%20&%20Eteokleous.pdf.

Mobile Learning. (2011). *Wikipedia*, Retrieved March 20, 2011 from http://en.wikipedia.org/wiki/Mobile_learning

Medical University of Tehran. (2011). Retrieved March 25, 2011 from http://www.tums.ac.ir/

Meeker, M. et al. (2009). Economy + Internet Trends," Morgan Stanley, Web 2.0 Summit. Retrieved March 25, 2011 from http://www.morganstanley.com/institutional/techresearch/pdfs/MS_Economy_Internet_Trends_102009_FINAL.pdf

Montazer, Gh. (2006). Chalesh-ha ve rahkarhai toseye etelaati nezame amozesh ali der Iran. In *Electronic Learning Conference,* Zanjan University, 24-25 May 2006. Zanjan: Zanjan University: 101-145.

Paczkowski, J. (2009). OMFG: 4.1 Billion Text Messages Sent Every Day in U.S. *Digital Daily*. Retrieved March 20, 2011 from http://digitaldaily.allthingsd.com/20091008/omfg-4-1-billion-text-messages-sent-every-day-in-us/

Peyam Noor University. (2011). Retrieved March 25, 2011 from http://www.pnu.ac.ir/

University of Hadith Sciences. (2011). Retrieved March 25, 2011 from http://vu.hadith.ac.ir/

University of Shiraz. (2011). Retrieved March 25, 2011 from http://www.shirazu.ac.ir/

The Impact of E-Journals on the Malaysian Scholarly Community

Hashim Roosfa

Centre for General Studies, The National University of Malaysia, 43600 UKM Bangi, Selangor, Malaysia. Email: roosfa@ukm.my

M.A. Mohd Yahya

Department of Media Studies, FASS, University of Malaya, 50603 Kuala Lumpur, Malaysia
Email: mohdyma@um.edu.my

Abstract: *Electronic journals are a new scenario for the Malaysian scholarly community. Thus, this paper presents the positive and negative impacts of the scenario. Currently, Malaysian universities are competing with each other to be in the list of top ranked world universities. To achieve the target, academicians are under pressure to become visible by publishing their works in journals indexed by various international databases. Hence, publishing research findings in e-journals has become compulsory among the Malaysian academics. The aim of this study is to identify the impact of electronic journals on the Malaysian scholarly community. Policies implemented by Malaysian universities pertaining to the rules for publishing works in electronic journals were studied. In addition, this study focuses on the impact of electronic journals on the academician career path, promotional exercise and rewards received by Malaysian scholars. The study includes input and responses by leaders of the Academician Association and directors and editors of university presses.*

Keywords: *Malaysia; scholarly publishing; electronic publishing; scholarly journal; academia.*

Introduction

Scholarly publishing in Malaysia has been undergoing rapid changes during the last two decades. The globalization of higher education and research institutions, university rankings and the emergence of electronic publishing especially for electronic scholarly journals are amongst the unprecedented changes that most universities and scholars in Malaysia have to face.

It has been 15 years since the first scholarly electronic journal was published in Malaysia. Ling, Mashkhuri & Pang (1997) pioneered the discussion with an overview of electronic publishing in Malaysia. In 1985, the first print version of *The Malaysian Journal of Computer Science* was published and 10 years later, in 1995, the first electronic version appeared on the Internet. Five years later, Roosfa (2000) recorded only six titles of electronic scholarly journals published by various institutions in Malaysia. During the 15-year period since 1995, the country experienced a remarkable increase in the number of scholarly electronic journals (Zainab, 1997a; Zainab, 1997b). To date, 25 Malaysian electronic scholarly journals are serving the world's scholars and researchers in various disciplines of knowledge. These electronic journals are among about 250 scholarly journals being published in the country. Hew (2001) focussed her discussion on the scholarly journal publishing in Malaysia.

Most of the scholarly journals in Malaysia are published by higher education institutions, especially by the public universities. With nearly 30 million population, the Malaysian government has set up new public universities and higher education institutions to serve the need for higher learning. MASTIC (2004) reported on the science and technology knowledge productivity in Malaysia in its bibliometric study.

In 2009, 20 public universities in Malaysia registered the enrollment of 121,323 students with 13.7% in various postgraduate programs. To fulfill the need of producing more professionals and scientists, the Malaysian universities are also increasing the number of students in 'critical' areas. There are

116,684 students enrolled in various programs such as pharmacy (2.6%), dentistry (1.4%), medicine (8.4%), law (8.7%), accountancy (19.8%) and engineering (59.1%).

The Higher Education Ministry is working seriously towards getting more international students onto the Malaysian university campuses. In 2009, the ministry recorded 22,389 international students enrolled with the biggest number of students from Iran (17.9%), Indonesia (16.5%), China (9.4%), Libya (5.3%) and Iraq (5.2%). In this circumstance, Malaysian public universities have to offer qualified and experienced lecturers conducting courses at various levels. In the same year, the ministry recorded 26,742 academic staff in public higher education institutions in Malaysia with 6,946 (26.0%) of them holding a doctorate qualification. Malaysian higher institutions are also served by foreign expatriates with the number of 1,428 out of the total number of academicians.

The Malaysian government is also targeting more international students to change the country's image into a higher learning hub for the region. To achieve this aspiration the facilities at the universities are being upgraded, especially the science laboratories; the number of research grants was increased, and the number of qualified lecturers is also being increased by focusing on increasing the number of lecturers with Ph.D. In addition, five public universities were upgraded to be research universities with one of the universities becoming an APEX university.

Research universities are expected to receive more research grants in the future years. Apart from that, the academicians in all these universities are being given more work and responsibilities. Nevertheless, their promotion and career path will depend on the outcome of their research, especially the number of articles published in journals listed in the international databases. Currently, only selected international databases are considered by the universities in the appraisal process of their academic staff. Therefore, the chances for the academicians to publish their works in journals listed in these databases are very slim especially for those in the area of arts, humanities and social sciences. Basically, this paper begins with a brief historical backdrop of scholarly journal publishing in Malaysia. This is followed by a discussion of the new trends in higher education in Malaysia. Emphasis is subsequently placed on the opinions of the actors among the prominent scholarly figures in Malaysia as well as the professionals of university presses in Malaysia. The paper closes with potential lessons for policy.

Historical Backdrop

The history of scholarly journal publishing in Malaysia started 165 years ago. *Journal of the Indian Archipelago and Eastern Asia* was the first scholarly journal published in Malaya (1847-1862). This was followed by *Journal of Eastern Asia* which appeared in 1875. Both journals were owned and published by individuals. The earliest journal published by a scholarly society was *The Journal of the Straits Branch of the Royal Asiatic Society* (*JSBRAS*) which came out in 1878. The journal then changed its title to *The Journal of Malaya Branch of the Royal Asiatic Society* (*JMBRAS*).

Basically, scholarly journals published in the early twentieth century aimed to serve the interests of the British colonial officers. The British set up many departments and research institutes, namely The Institute of Medical Research, Rubber Research Institute, Agriculture Department, Forestry Department and the States Museum. Among the journals published during that period were the *Agricultural Bulletin of the Malay Peninsula* (1891), *Journal of the Straits Medical Association* (1892), *Perak Museum Notes* (1893), *Sarawak Museum Journal* (1911), *The Planter* (1920), *Medical Bulletin* (1923) and *Quarterly Journal of the Rubber Research Institute of Malaya* (1929).

The first scholarly journal published by a university was *Jernel Fakulti Kejuruteraan Universiti Malaya*. It was a journal published by the Engineering Faculty in the University of Malaya in 1958. Four years later, in 1962, the journal changed its title to *Journal of Department of Engineering University of Malaya.* Lim (1975) elaborated on the problems of publishing scholarly journals in Malaysia. Stone (1975) elaborated on the scholarly serial publications of academic institutions and societies in Malaysia during 1970s.

To accomplish the need for higher education, the government developed more universities. Much research on documentation of Malaysian scholarly journals has been conducted since the middle of the twentieth century, such as Tiew (1998, 1999 and 2003). *JMBRAS* remains as the longest surviving scholarly journal in Malaysia, from 1878 until now.

Currently there are about 250 Malaysian journals on the market, facing many challenges and problems. Lack of professionalism among the editors, poor refereeing systems, bureaucracy and financial problems and too many general and multidisciplinary journals competing to get articles from the same group of authors are among the common problems faced by most of the journals. Hence, these problems result in the shortage of good articles published. In addition, low quality and lateness of publication of up to 2-3 years are a common scenario for some Malaysian journals. The majority of the Malaysian academicians are bilingual because the Malaysian education system provides them with communication skills in Malay Language and English. If they decide to contribute their articles only to international journals abroad, this will affect the local journals badly.

The attitude among the majority of Malaysian scholars, who now publish their work in e-journals abroad, has had this effect. Presently, Malaysian scholars' performance is rated according to their articles published in e-journals, especially those listed in the ISI or Scopus databases. They have to send their articles to journals published abroad because there are only thirty-five Malaysian journals listed in Scopus and five in ISI. Consequently, what will happen to local scholarly journals in Malaysia? To date, there are about 250 scholarly journals on the market. The move to encourage Malaysian scholars to publish their works in journals listed in international databases such as ISI and Scopus is a positive effort in increasing the visibility of Malaysian authors. Nevertheless, there should be a mechanism for sustaining and positioning senior Malaysian scholars, especially in arts, social sciences and humanities in their academic advancement. Other international journal databases are even better in terms of maturity and subject area coverage. This paper will also study the impact of publishing in international electronic journals on scholarly book publishing. Directors and editors of university press were contacted to give their views.

The Impact of Electronic Journals

The importance of diversifying into electronic media is still a hot topic among the Malaysian academicians. Despite the increasing number of Malaysian works published and listed in internationally indexed journals, there is always a negative side of the scenario. Harris (2008) discussed the ranking of journals.

Positive Impacts

University of Malaya (UM) acknowledges the contributions of its academic staff who publish their work in ISI journals. The university is giving incentives to academics for their publications that are indexed within the year. Hence, the University of Malaya has outlined its reward incentive policy: except for single authorship, claims for multiple UM authors must be initiated by one of the authors and must be approved by the other authors. For publications with more than one UM author, any author can make the claim, but approvals must be obtained from all co-authors. Amounts to be apportioned between the co-authors will be decided among the authors themselves. Each author can only claim a maximum amount of RM50,000 in a publication year.

Impact Among Academicians – Looking Abroad

Searching, quoting, citing and referencing fellow researchers' works became easier with e-journals. E-journals enable the researchers, including the academicians, and their research work to be visible to other researchers globally. In addition, e-journals listed in the international databases contribute to the better rankings of the university. These are among the benefits of publishing in e-journals.

In a recent development, Times Higher Education Supplement (THES) only uses ISI data in its university rankings mechanism. University of Malaya followed this step by only taking into consideration its academicians' works published in the ISI database. The works published in journals listed in other databases including Scopus carry no weight. Recent developments prove that University of Malaya's efforts have yielded positive results although this was implemented only a few years ago. The university was listed among the top 200 universities in the THES 2009 list. On the other hand, Malaysian research universities are trying to obtain the highest number of records in ISI Web of Science (ISI WoS). This healthy competition is presented in Tables 1-4.

Table 1. Malaysian published works in ISI WoS and Scopus by Malaysian research universities, 2010

Universities	ISI WoS	Scopus
UM	361	389
USM	366	428
UKM	166	229
UPM	216	333
UTM	85	130

Note: UM = University of Malaya, USM = Universiti Sains Malaysia, UKM = Universiti Kebangsaan Malaysia, UPM = Universiti Putra Malaysia, UTM = Universiti Teknologi Malaysia

Table 2. Citations/record of Malaysian universities in ISI WoS and Scopus, 2005-2009

Universities	Number of records		Citations		Citations/record	
	ISI	Scopus	ISI	Scopus	ISI	Scopus
UM	3840	4693	6315	7506	1.64	1.60
USM	3612	4537	6914	8755	1.91	1.93
UKM	2608	4048	3004	4516	1.15	1.12
UPM	2612	4124	3359	4763	1.29	1.15
UTM	1330	2092	1494	2344	1.12	1.12

Table 3. Ratio of citations/academician among Malaysian universities ISI WoS and Scopus

Universities	Number of citations 2005-2009		Number of academicians 2008*	Citations/academician 2005-2009	
	ISI	Scopus		ISI	Scopus
UM*	6315	7506	2077	3.04	3.61
USM*	6914	8755	1780	3.88	4.92
UKM**	3004	4516	1563	1.92	2.89
UPM*	3359	4763	1938	1.73	2.46
UTM*	1494	2344	1905	0.78	1.23

Table 4. Ratio of records/academicians among Malaysian universities

Universities	# of records 2005-09		Academicians 2008	Records/academics 2005-09	
	ISI	Scopus		ISI	Scopus
UM*	3840	4693	2077	1.85	2.26
USM*	3612	4537	1780	2.03	2.55
UKM**	2608	4048	1563	1.67	2.59
UPM*	2612	4124	1938	1.35	2.13
UTM*	1330	2092	1905	0.70	1.10

Source: *Statistics from Planning and Research Division, the Ministry of Higher Education Malaysia
** Malaysian Research University data 2010.

Impact on Local Journals – More Visibility

The electronic journal has changed the direction of many journal publishers and editors in Malaysia. Currently, they are working more towards having their journals listed in international index services. In addition, the Ministry of Higher Education is monitoring the status of Malaysian universities in the world university rankings. Due to that, Malaysian universities are concentrating their effort on increasing their academicians' visibility in international indexing and abstracting services. Since

THES is using ISI's Web of Science (WoS) in measuring a university's citation performance, more Malaysian journals are now focusing on publishing in journals whose contents are listed in indexing services such as WoS. To date, there are 11 Malaysian journals listed in ISI WoS and another 44 in Scopus. Although the total number of Malaysian journals in ISI and Scopus is small compared to the total number of journals published in Malaysia, the figure is increasing steadily year by year. Nevertheless, many local journals are facing difficulties in getting good manuscripts from renowned authors. At the same time, the number of titles in terms of books and monographs published by university presses in Malaysia is dropping tremendously as confirmed by Salleh Yaapar, the Director of USM Press and Abdul Manaf Saad, the Director of UM Press. Furthermore, Saadah Jaafar, a senior editor (social sciences and humanities) at UKM Press expressed her concern at the latest development where authors are more interested in revising and publishing their theses or dissertations as journal papers than turning them into books or monographs.

Impact on the Career Path

As noted, Malaysian universities and the Ministry of Higher Education are encouraging university lecturers to publish their research work in e-journals. The number and the impact factor of articles published in Scopus and ISI influence the career path of Malaysian scholars especially in the process of recruitment of new academicians and in the promotion exercise. The promotion practice and the recruitment process of a new academician are also influenced by their publications in e-journals, especially those journals listed in the Scopus or ISI databases. The reason behind this policy is mainly that ISI WoS database is being used by the THES in its university rankings process.

In the University of Malaya, a candidate for the position of university lecturer must not only hold a doctorate degree but must also have published his or her work in a journal listed in the ISI database. Academicians of research universities in Malaysia also need to publish their works in ISI journals for their promotion to higher positions. But, compared to University Malaya, other research universities accept works published in journals listed in Scopus as well as ISI. The move was aimed at increasing the visibility of Malaysian scholars. The example of promotion criteria for UM staff is presented in Table 5. Books accepted for promotion are only those published by an established international publisher or a recognized local or international university press including UM Press.

Table 5. Promotion criteria for non-science disciplines at University of Malaya

# of publications required for promotion	**Books**	**# of ISI / Refereed Articles in Top Tier Journals (Category A & B) / Chapters In Books**
Professor A	6	**50**; at least **16** must be in Tier 1 & 2 ISI Journal
Professor B	4	**40;** at least **12** must be in Tier 1 & 2 ISI Journal
Professor C	2	**30;** at least **8** must be in an ISI Journal
Associate Professor	1	**15;** at least **4** must be in an ISI Journal
Senior Lecturer	-	**10; a**t least **2** must be in an ISI Journal
Lecturer (for confirmation)	-	**5;** At least **1** must be in an ISI Journal after appointment to Lecturer post

Note: 1. Category A – Top 50 journals to be used for the whole faculty/academy
2. Category B – Top 50 journals to be used for each department in the faculty/academy

Brain Drain

The significant effect of the move is 'brain drain' among the research university senior lecturers and professors. They have published long lists of papers and articles in various journals in their field and this is the time for them to sit down and write books. Nevertheless, with the new university regulations, they still need to publish in high impact journals. To achieve the Key Performance Index (KPI), they must publish at least three papers in any high impact journal per year. Some senior professors have left the university to join new universities and some have even ended by doing nothing at home.

Due to the high expectation, there are cases where a candidate offered a junior position at University Malaya rejected the offer to accept a better offer at a new university. Senior professors who failed to achieve such KPI have to decide whether to accept a lower contract position or to end their career at the university. In the Malaysian service system, there are three levels of professor with grade A, B and C, where the most senior professor has professor A status. There are cases whereby a professor at the 'B' grade who failed to achieve the KPI accepted an offer to be in grade 'C' and in other cases a professor at 'C' grade rejected the offer to be in a lower position as an associate professor. In other scenarios, lecturers left research universities to join new universities to avoid the heavy burden of the working environment in research universities.

Associate Professor Aziz Abdul Rahman, Secretary General of the Academic Staff Association, University of Malaya, believes that the moving of the university academic staff from University of Malaya to various new universities is influenced by pull and push factors. Professors are terminated based on the fact that they failed to achieve the university's Key Performance Index, which requires them to publish their works in high impact indexed journals. At the same time, the professors were offered the same position at new universities. Both parties, the professors and the new universities, benefitted. To the association, this scenario cannot be looked at as 'brain drain' since the professors are still delivering their services at local universities.

Table 6. University of Malaya professors terminated (April 2011)

Academic Unit	N
Faculty of Arts and Social Sciences	7
Faculty of Economics	2
Faculty of Education	1
Cultural Centre	1
Academy of Malay Studies	4

Conclusion

In conclusion, this study revealed the implications of electronic journals for academic enhancement in Malaysia. Undoubtedly, Malaysian universities, especially the research universities, are becoming more visible as their academic staff contribute to the increasing number of articles published by international journals which are indexed by international index services such as ISI WoS and Scopus.

Nevertheless, the requirement to publish in international electronic journals has become a big challenge or an obstacle for senior academicians to advance in their academic career. Besides, the exercise discourages students from joining the academic line due to the high expectations and publication policies. Consequently, some of the senior professors left the universities to join new universities as they could not achieve the requirement or targeted performing index set by the university.

References

Harris, C. (2008). Ranking the management journals. *Journal of Scholarly Publishing,* 39(4): 373-409.

Lim, H.T. (1975). Problems of publishing scholarly journals in Malaysia. In Beda Lim (Ed.)*Scholarly publishing in Southeast Asia* (pp. 57-69). Kuala Lumpur: ASAIHL.

Ling, T.C., Mashkhuri Yaacob & Pang, K.K. (1997). An overview of electronic publishing. *Malaysian Journal of Library & Information Science*, 1(2): 1-11.

Hew, M. (2001). Scholarly journal publishing in Malaysia. *Journal of Scholarly Publishing,* 32(3):164-168.

MASTIC. (2004). Science and technology knowledge productivity in Malaysia bibliometric study 2003. Putrajaya: MASTIC.

Roosfa, H. (2000). Scholarly E-journal in Malaysia: the status and academician attitudes. Dissertation MLIS. Faculty of Computer and Information Science, University Malaya.

Roosfa, H. (2002). The development of science, technology and medical scholarly journals in Malaysia, 1890-2002. Paper presented at The 11th International Conference for Science Editors, 24-28 August 2002, Beijing. China.

Stone, B.C. (1975). Scholarly serial publications of academic institutions and societies in Malaysia to-day: A review and commentary. In Beda Lim (Ed.) *Scholarly publishing in Southeast Asia* (pp. 38-45). Kuala Lumpur: ASAIHL.

Tiew, W.S. (1998). History of *Journal of the Malaysian Branch of the Royal Asiatic Society (JMBRAS)* 1878-1997: an overview. *Malaysian Journal of Library & Information Science*, 3(1): 43-60.

Tiew, W.S. (1999). Some scholarly English periodicals in pre-independent Malaysia: an historical overview. *Malaysian Journal of Library & Information Science,* 4(1): 27-40.

Tiew, W.S. (2003). Malaysian historical journals (1847-1966): A bibliometric study. *Malaysian Journal of Library & Information Science,* 8(1):19-43.

Zainab, A.N. (1997a). Malaysian scholarly journals: Their visibility and characteristics. Dlm. Md. Sidin Ahmad Ishak (Penyelenggara). *Penerbitan jurnal ilmiah di Malaysia*. Kuala Lumpur: Penerbit Universiti Malaya.

Zainab, A.N. (1997b). The coverage of Malaysian scholarly journals by international indexing and abstracting services: an assessment of their visibility and common characteristics. *Malaysian Journal of Library & Information Science*, 2(1):13-30.

Workhops

Hacking Scholarly Communication and the Scientific Record

Leslie Chan, with contributions from Peter Pennefather, Gale Moore, Frances Garrett, Matt Price, Jason Boyd, and Sally-Beth MacLean

University of Toronto Scarborough, 1265 Military Trail, Toronto, Ontario, Canada.
Email: chan@utsc.utoronto.ca

Abstract: *Current scholarly communication practices are highly tethered by publishers' interests and constrained by practices from the print-based era. The workshop will introduce and explore ways by which new scholarly practices can be "mobilized" so that the key functions of scholarly communication - certification, quality control, archiving, and rewarding - can be decoupled and better served by emerging tools for collaborative and open authoring, sharing, and reputation management.*

At the same time technology now exists for allowing authors to launch, enable access and curate data streams, turning their analytical work into collaborative spaces. This untethering process could facilitate others in drawing their own insight from the original scholarly work applied to formatting those data streams and create new opportunities for scholarship as well as accommodating retention of original authorship information. This also enables the idea of an inside-out, data driven, publishing format where publications are embedded in the data stream record rather than the other way around.

Purpose: The intent of the workshop is to engage participants in discussion and thinking about new and potential developments in scholarly communication and how new tools could be developed and implemented collaboratively. The outcome of the workshop will have important implications for the future of ELPUB, as the conference series continues to define its place in the growing international discussion on the future of scholarly communication, and about the nature of the changing medium, practices, recognition of authors' contributions to the conference series and how best to improve the visibility of ELPUB.

Ideas to consider:
Document formats: ePUB, PDF, XML etc. (interoperability and mobility)
Workflow system: peer review, quality control etc
Publishing platform: wiki, blog, WordPress, etc.
Data streaming: various data formatting standards that enable extensive metadata annotation, TIFF, PDF, etc
Enhancement: Tagging, open bibliography,
Metrics: article level, download, Google analytics, citation, etc.
Putting them all together - support, infrastructure, etc.

Structure: A wiki will be set up prior to the workshop and participants input will be sought in advance. The workshop will consist of breakout sessions, followed by collaborative conceptualization of what a viable, mobile, and scholar driven communication system may look like. Participants will gain sufficient knowledge and resources to further explore possibilities and to join in the appropriate communities or networks for development.

Keywords: *Scholarly communication; new scholarly practices.*

Applications for Efficient Use of E-content

Tayfun Başal

Elsevier Country Manager, Turkey and Central Asia. Email: T.Basal@elsevier.com

Tools that Support Search Platforms: The Road Map of Applications for Discovery of Science

Kıvanç Çınar

Director, Informascope, Turkey. Email: kcinar@informascope.com

Snuggled with an E-book (Reader): How to Increase Use of your Library's E-book Collections?

Uygar Saral

Founder, Selekt Computers, Turkey. Email: usaral@reeder.com.tr

E-book Devices and Respective DRM (Digital Rights Management) Solutions

Digital Publishing Solution for eBooks

Orçun Madran

Atılım University, Turkey. Email: omadran@atilim.edu.tr

Bora Ünal

Atlantik LTD, Turkey. Email: bora@abi.com.tr

Abstract: *The demand for eBooks is accelerating rapidly — the percentage of book sales tripled in the first five months of 2010 (According to the Association of American Publishers, eBooks made up 8.5 percent of trade book sales in the first five months of 2010, compared to 2.9 percent in the same period in 2009). Book publishers, distributors, retailers, and consumers eager to seize this opportunity, Digital Publishing offers a comprehensive software solution for authoring, securely distributing, and reading digital books on computers and a wide range of mobile reading devices.*

Digital Rights Management (DRM) is a target solution for content providers and publishers to make sure their intellectual properties remain undamaged. DRM may applied to content several ways. The main question does it provide a good user experience?

Main topics of the Workshop are as follows:

- *Authoring: Cost-effective workflows.*
- *Content Protection: Secure content distribution*
- *Delivery: Flexible cross-device purchase and consumption.*
- *What others are doing besides DRM?*

Keywords: *Digital publishing; eBooks; Digital Rights Management.*

Semantic Web

Ana Alice Baptista

University of Minho, Portugal. Email: analice@dsi.uminho.pt

Silvana Vidotti

UNESP, Brazil. Email: vidotti@marilia.unesp.br

Abstract: *The Semantic Web (SW) workshop will focus on a practical view of the field. It will start with an overview of the SW fundamentals: the layercake model, the principles behind the Resource Description Framework (RDF) and the linked data approach. It will, then, concentrate on the interoperability of metadata, taking into account properties, values and the use of encoding schemas. The need for machine-readable values (objects) and the way to do it will be highlighted. Practical examples will be given throughout the workshop.*

Keywords: *Semantic web; Resource Description Framework.*

Author Index