INFuture2013

The Future of Information Sciences

Information Governance

Department of Information and Communication Sciences
Faculty of Humanities and Social Sciences
University of Zagreb

THE FUTURE OF INFORMATION SCIENCES

INFuture 2013

4th International Conference
"The Future of Information Sciences:
INFuture2013 – Information Governance"
Zagreb, 6-8 November 2013

Organizer
Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences, University of Zagreb
under the high auspices of the President of the Republic of Croatia
supported by the Mayor of the City of Zagreb

Technical editor
Hrvoje Stančić

Cover design by
Hrvoje Stančić

Publisher
Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia

Printed by
Inter-ing, Zagreb

Impression
150 copies

All published papers undergo an international, double blind peer review.

ISSN 1847-8220

THE FUTURE OF INFORMATION SCIENCES

INFUTURE2013

INFORMATION GOVERNANCE

Edited by

Anne Gilliland, Sue McKemmish, Hrvoje Stančić,
Sanja Seljan, Jadranka Lasić-Lazić

Zagreb, November 2013

CONTENTS

Preface

This publication is published as a result of the fourth biannual conference *The Future of Information Sciences* – INFuture 2013: Information Governance, organized by the Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences, University of Zagreb and this time, under the high auspices of the President of the Republic of Croatia and with the support of the Mayor of the City of Zagreb.

Information sciences, as an interdisciplinary field, have developed at the intersection of computer sciences, information processing, technology, philosophy and organization, respecting scientific findings, practical implementations, management logic, computer formalities, market and industry demands.

Information governance, as part of information sciences, focuses on the right information in the right context, logically interconnected, stored, retrieved and implemented, but is also realised through the strategic aspect in the modern archival practice, bringing various benefits as well as risks.

Results of the research in the section *keynote papers* address broad issues of global digital infrastructure for archival access supported by digital curation and semantic web, cloud computing environment and related recordkeeping and archiving benefits and risks as well as connection and influence between machine translation and translation studies.

The section *theory and methodology of information governance* deals with the domain specific records management, international standard descriptions, best practices, human role in project management, digital records in an increasingly networked society, various domain-specific information systems.

The topic *knowledge management* elaborates on classification system using public opinion, use of social services and their side effects and social lean thinking through social business environment.

Various *applications* have been presented *for e-society and e-government* ranging from remote access to statistical microdata, electronic signatures, digital rights management in the mobile application market and music industry.

The topic of *language technologies* is oriented to terminology extraction, use of machine translation systems and cross-language information retrieval.

The topic of *digital curation* deals with crowdsourcing the digital cultural heritage in museums and galleries and digital curation and preservation of the online exhibits.

In the section related to *new challenges in interdisciplinary education*, several researches have been presented relating to computer-based assistive technologies in education for students with disabilities, presentation of social sciences and humanities in media, map of various competences, lifespan of web references, role of information technology in the social role of film, interactive ap-

plication for Latin language learning, and aspect of e-learning and ICT education for elderly.

The presented topics reveal interest for applications used by broad population, for interdisciplinary education, strategic harmonization with the European policies, business ethics and national goals, for integration into multilingual society preserving the heritage properly, for knowledge organization and information management. We believe that the results of the research will help us find new ways of building information resources and e-services which would contribute to better information governance in the world overcrowded with information.

Editorial Board

KEYNOTE PAPERS

'To the Tasks and the Skills'*: Considerations and Competencies for Designing Glocal Archival Access Systems

Anne Gilliland
Center for Information as Evidence
Graduate School of Education and Information Studies
University of California Los Angeles (UCLA), USA
gilliland@gseis.ucla.edu

Summary

This paper lays out multiple roles that archives and their contents currently play within and across different jurisdictions and communities. It argues that digital curation and the Semantic Web respectively offer the archives field a model and a structure for developing a global digital infrastructure for archival access that also might fit within broader digital description and access activities. However, as currently expressed neither provides sufficient support for the complexities of designing archival access systems and services that can simultaneously address local, national and transnational considerations. The paper proposes a set of principles, with examples of approaches, that could be applied in designing archival infrastructure, identifying necessary tasks, and nurturing competencies that promote accountable and equitable ways of addressing diverse communities' access needs and uses for archival resources in a glocal world.

Key words: Archives, Digital Curation, Semantic Web

Introduction

Current discourses within and across the archival profession, information technology and management communities, and disciplinary scholarship locate the archival role in varying places along a wide-ranging spectrum. That spectrum encompasses not only the support of scholarship and more general historical uses, but also of bureaucratic recordkeeping, governance and regulation; accountability, transparency and legitimation; long-term research data curation;

* How long will it be 'till we've turned
To the tasks and the skills
That we'll have to have learned
If we're going to find our place in the future
And have something to offer
Where this planet's concerned?
Jackson Browne, "How Long," *World in Motion*, 1989.

cultural heritage management; empowerment and representation for marginalized communities; and human rights concerns. This diversity of roles, and how they are reflected in the nature, scope and use of archival holdings, underscore the richness that the archival dimension brings to discussions about the ways in which information access infrastructure should be conceived and implemented within the archival field as well as when interfacing or integrating with other information domains.

Such richness demands complexification, heterogeneity and hierarchy rather than the tendency towards simplification, homogeneity and “flat thinking” that often characterises how we approach the building of information access infrastructures. It begs the question, *can a wide-scale archival access infrastructure support complexity, heterogeneity and hierarchy in efficient and distributed ways?* Moreover, these different archival roles can find themselves at odds with each other, especially within individual institutional missions and governance frameworks, and no single archives is likely to be able to address all roles to the same extent. *How then might a more systematically-developed archival access infrastructure help to mitigate these limitations?*

Further adding to this complexity, in ways that are both generative and confounding, are the demands of and interaction effects between two aspects that archives have often been criticised for marginalizing or simply ignoring. The first of these are the needs, modalities and vocabularies of local grass-roots, identity- and issue-based, and activist communities that may feature as the subjects or objects of archival materials, especially those created under oppressive and colonial regimes, but that rely upon those materials for political (e.g., redress and reparations), social, cultural and personal purposes (Flinn 2011). *How might existing archival access infrastructures be reoriented to address the very specific and immediate needs of overlooked, underempowered, oppressed or recovering communities?*

The second aspect is the relative lack of contemplation on the part of individual archives, national and regional archival systems, and archival endeavours in the developed world in general, of the new needs and considerations arising when disseminating metadata and digital content online to a diversity of cultures, political regimes and communities in less advanced stages of development around the globe. *How might archival description and access practices be re-thought to address the implications of how archival materials are being described, disseminated and used in a globally networked world where user communities may be technologically or educationally disadvantaged; politically constrained; or invisible, widely distributed, or poorly understood because of distance, diaspora, or difference?*

There have been increasing calls from concerned communities, both local and those distributed around the globe, as well as from within the archival profession itself to engage in reconciliatory, culturally-sensitive and socially responsible ways to support these plural archival needs. These calls have taken many

forms, including Indigenous protocols for managing, disseminating and using archival materials (ATSILIRN 2012; First Archivist Circle 2007); professional statements of archival values (Society of American Archivists 2012); critiques by archival studies scholars (Jimerson 2009; Harris 2007); and challenges to existing professional education (Gilliland 2011). The concept of the Archival Multiverse in particular has been gathering momentum as a platform for pluralizing how archivists approach their work and are professionally educated (PACG 2011). The Archival Multiverse,

> includes a recognition of different ways of knowing, evidence paradigms, forms of archives (such as Indigenous keeping places) and records, and transmission methods; and acceptance of differing constructs of ownership, intellectual property, privacy, access, and rights in records, as well as what constitutes the secret and sacred material of different cultures in different space-times. Respectful, negotiated community partnerships, sharing of governance and decision-making about current and historical recordkeeping, and implementation of community-centric protocols such as those emerging from Indigenous communities are essential features of such approaches. As is the use of digital and social media technologies to support efforts to pluralize recordkeeping and the contents, uses, and development of archives, as well as digital repatriation (Gilliland, Lau and McKemmish 2013).

Somehow archival access systems must be able simultaneously to address both of local and global aspects. Tying the two strands together, we could ask the following: *What might an archival access infrastructure look like that was explicitly designed to be reconciliatory, culturally-sensitive and socially responsible, especially one that was world-wide in conception and operation?* The answer, I would posit, lies in the development of a multi-level *glocal* archival access infrastructure (Robertson n.d.; Robertson 1997). While national, sector, and institutional interests and perspectives often set the agenda in the past, today there is a growing focus on simultaneously addressing local and global needs and perspectives, as well as on understanding the effects of interaction between the two, i.e., the phenomenon known as *glocalization*. I have argued elsewhere that,

> Governments, business, research and socialization are increasingly conducted on a transnational and trans-community basis using the web and cloud technologies. Indigenous, local, and colonial recordkeeping and memory systems and practices that developed out of particular national and cultural traditions and worldviews are simultaneously grappling with their tangled and often still traumatic histories, and interfacing and negotiating with other traditions and worldviews in whole new digital ways. New and reconstituted nations are working to develop their national identities and strengthen their own bureaucratic, scholarly and profes-

sional infrastructures internally, while seeking visibility and recognition on the world stage (Gilliland 2013 in press).

These tendencies are directly relevant to the multiple archival roles and concerns laid out above.

Two current developments that are wider than the archival field hold out the potential to help in the development of such an archival access infrastructure--digital curation offers the archival field a procedural and collaborative framework within which these activities can be situated, and the Semantic Web offers a metadata infrastructure for sharing data and embedding exploitable relationships via the Web between archival holdings and also other digital information resources around the globe. However, there are significant metadata and retrieval aspects that need to be more fully articulated and their feasibility investigated in order to provide conceptual and practical support for the complexities of designing archival access systems and services that can simultaneously address local, national and transnational considerations. This paper proposes some principles that might guide this reorientation of archival access systems and an overarching archival access infrastructure, and provides some examples of the tasks and skills or competencies that will be necessary to address them.

Background

Historically, archives, especially institutional archives, lacked imperatives and incentives to cooperate and functioned in relative isolation. They were isolated from each other, from the creators of the records that they would eventually acquire, and from other institutional players in the information ecosystem such as libraries and museums. Moreover, archival traditions (i.e., the conceptual and practical bases upon which archivists were trained and repositories operated) differed in significant ways from region to region around the world, an aspect that remains under-accounted for in archival descriptive standards. Some key events from the early 1980s onwards began to challenge the reasoning and technical practices behind such isolation. The development of standardized machine-readable description such as the MARC Archival and Manuscripts Control (AMC) format in 1984 facilitated the emergence of shared repositories of archival descriptive information such as the Research Library Information Network (RLIN), as well as the integration of those descriptions into institutional library OPACS. The development of Encoded Archival Description (EAD) in the mid-1990s provided a more “archival” infrastructure capable of providing serving as a backbone for web-access to online digitized archival materials. While these developments have certainly been capable of supporting global access, they nevertheless emanated out of ideas and practices that were oriented around the interests of the parent institution as well as historical scholarship, the prevailing archival tradition, and dominant cultural constructions in the originating jurisdiction and thus did not meet emerging glocal concerns.

In the 1990s, new archival conceptualizations began to significantly challenge existing archival traditions. The articulation of the Australian records continuum model in 1996 highlighted the societal and structural inadequacies of both the life cycle and exclusively custodial approaches to archiving in which archivists were divorced from the processes of records creation and only became involved with them after they were no longer being actively used by the creator (Upward, McKemmish and Reed 2011; Upward 2006 & 2007). In concert with this latter point, electronic records archivists in the same decade urged archivists to intervene at the point of records creation to ensure the creation and preservability of a reliable and authentic digital record. Although archival theory had always stressed the organic nature of records creation and use, these new conceptualizations underscored that creation, preservation, access and use are all closely inter-related and inter-dependent.

This might all seem like old history today, given that archival descriptive metadata and digitized copies of archival material are being virtually collated in shared or federated regional and disciplinary repositories and information systems; that the metadata is being widely exposed for harvesting and searching; and that the assignment of linked open data by archivists is becoming more common. However, regardless of the size and scope of these initiatives, this trans-institutional and Web-oriented movement remains something of a patchwork quilt of possibilities and projects that continues to represent and favour the perspectives and priorities of individual institutions and funders. Metadata standards and their origins can be similarly critiqued. As for information retrieval, the focus of so much work over the years in library and information science, this aspect, and the particular challenges and opportunities that archival information retrieval might present, have barely received any attention at all.

New considerations are pushing the field to think more strategically and collaboratively about archival access infrastructure, however. For example, Cloud storage and access have highlighted questions about jurisdiction, management and control over digital data, electronic records and cultural information resources that have been collaboratively created and used over networks. Cooperative and collective digitization initiatives have increased the amount of digital content available over the web to potential user communities around the world but, as already discussed, have raised questions about whether one size fits all in terms of how that content should be described or might be searched or disseminated. Digital repatriation offers ways for often unique materials to be in multiple locations at once but also offers the prospect of developing multiple and community-sensitive local management, metadata and access regimes for the same content.

Proposed principles

The following are several proposed principles to help in the identification of tasks, competencies and actors necessary for the reorientation of archival access

systems. These principles reassert much of what has already been said in the critiques and debates discussed above, but are applied specifically to archival access:

- *Archival access systems cannot stand apart from other information infrastructure:* Archives are only one component in an organic world of information creation, management, use and reuse. Archival descriptive and retrieval mechanisms must be able to interface with the rest of that world, especially as understandings of what constitute archival as opposed to other kinds of information objects become increasingly blurred.
- *Archival access and other archival and recordkeeping functions are interdependent:* Access as a function cannot be considered in isolation from the networked ways in which records and other archival materials are digitally created, managed and used.
- *Rich context is essential*: The need for context becomes increasingly important as the volume of archival content online multiplies and becomes accessible at more and more granular levels, and as that content is read both “along” and “against” the grain by diverse users. Context needs to be further captured and exploited, whether by manual or automatic means, by archival institutions or crowd-sourced, or by professional or community experts. Context, however, should not be completely hard-coded—to facilitate fluidity in interpretation across communities and over time, and to support a plurality of narratives and counter-narratives, mechanisms must be in place to facilitate the creation and maintenance of multiple simultaneous descriptions, as well as periodic re-description of the same content (in other words, just as archival materials take on different meanings and are subject to different interpretations as they move through time, description should not be a one-time or one-agent activity). It should also be possible for users to contextualize and recontextualize archival content according to their own needs, perspectives and modes of expression (e.g., by developing or identifying alternative or counter-narratives) and contribute this contextualization back into the system.
- *Promoting and exploiting differing levels of granularity as well as existing hierarchical relationships between different metadata schemes and approaches will introduce heterogeneity into how archival content is described and made available for retrieval while reducing metadata overhead*. Traditional archival description principles uphold the use of natural hierarchies that exist within archival collections and also the tenet that all content does not have to be described to the same level of detail. Given that different descriptive metadata schemes can today be applied that operate at every level from the item (indeed even within-item) level all the way up to the most general description of an archival collection or aggregation (in fact, to the level of the repository itself), the requirement for every archival institution to work at multiple levels is somewhat allevi-

ated. Instead it should be possible to make strategic decisions about when there might be a compelling need to do rich descriptions, potentially addressing specific identified needs of particular communities (for example through the use of pluralized access points, complex authority files that address co-creator roles, and bilingual descriptions), and when a higher-level approach might suffice.

- *Pluralism and complexity are defining, and arguably the most emancipatory, characteristics of the Archival and indeed the wider Information Multiverse.* In devising standards, best practices, regulations, and terminology for international implementation, archival developers need to focus less on getting everyone to do things the same way, and more on how to inter-relate diverse community practices and ontologies. This will support more equitable, consultative, and culturally sensitive exchange of metadata and content across national, cultural, linguistic, and ideological boundaries, as well as among different institutional settings and technological environments.
- *Power differentials and inequities are at work in every aspect of information creation, preservation, organization, access and use.* The development of access infrastructure, including archival descriptive metadata and information retrieval must directly acknowledge and address the negative consequences of such power dimensions, and actively support emancipatory practices.

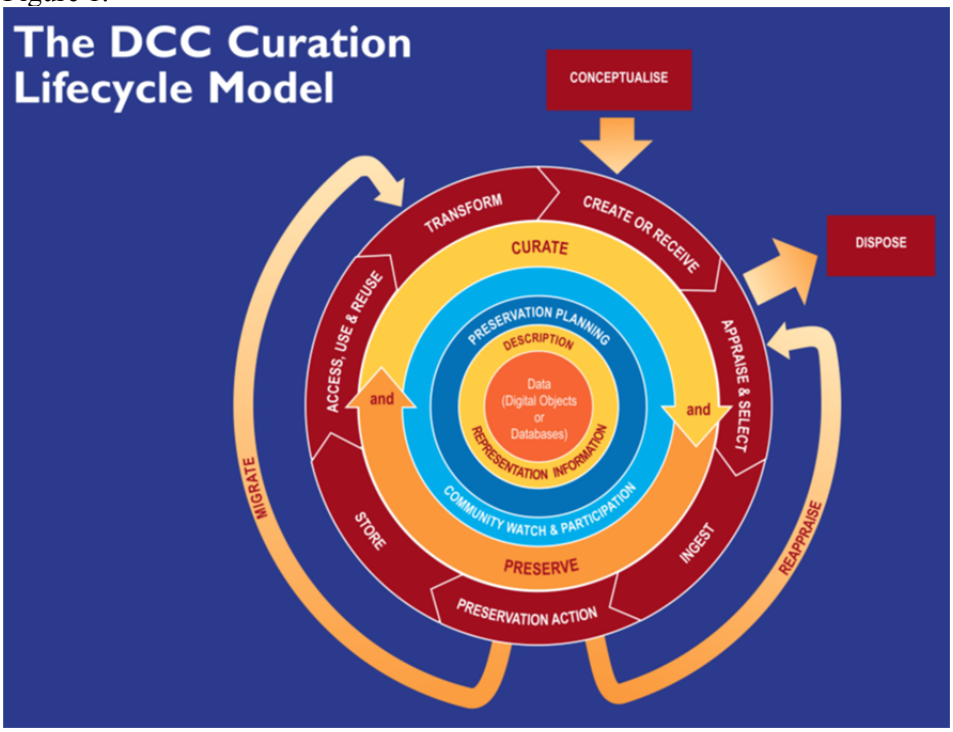
Applicability of the Curation Life Cycle Model and the Semantic Web

With these principles in mind, we should first consider whether we already have models and structures upon which we can build that can engage in participative and consultative approaches and also capture the nuances, contingencies, uncertainties and semantic and linguistic diversity, archaisms and semantic drift that are integral to working with archival content.

The term "digital curation" has emerged in recent years as a transprofessional umbrella covering strategies, technological approaches and activities involved in creating, managing, preserving, disseminating, and using digital content (DigCCurr n.d.; Higgins 2011). The Digital Curation Center (DCC)'s Curation Life Cycle Model (see Figure 1) can be differentiated from the traditional archival life cycle approach by its emphasis on multiplicity, dynamism and awareness and engagement of all the parties that might be involved, from creation all the way to consumption. Although the model retains some of the linearity that the Records Continuum Model rejects, access is clearly related to and interdependent upon the other processes depicted. Many different curatorial activities, with many different associated agents, are occurring simultaneously Content is seen to be ever evolving in value, and being continuously transformed through use into new content. Metadata (i.e., Description and Representation Information) is explicitly identified as being central and integrally re-

lated to both the content and the encircling processes. In conception, therefore, it would seem that this model is very congruent with the dynamism and multi-agency aspects laid out in the principles above. Although many digital repositories conduct their digital curatorial activities with a single Designated Community (to use OAIS parlance) in mind, *what if we were to create an overlay for this model that, for each facet, delineated glocal, pluralized considerations?* Such an overlay could be used as a tool to raise awareness when preparing future archival professionals, and also by standards developers, archival institutions, and archival professionals to identify policies, metadata practices and technological solutions that might enfranchise more communities and uses.

Figure 1.



Source: The Digital Curation Centre

Although archives have been somewhat slower than many other information communities to implement RDF in their descriptive practices, the Semantic Web and its linked data approach does provide an infrastructure that supports bridging between different metadata schemes and disparate materials. However, vagueness, uncertainty and inconsistency have been among the criticisms leveled at the effectiveness of the Semantic Web in delivering on its promise. In the archival world, and when thinking about the complexities of a glocal and pluralized approach, epistemic and ontological “messiness” is endemic, seman-

tic drift is inevitable as historical materials move through time, much about archival materials may be unknown or incomplete, the same materials may be viewed from completely different perspectives, and hierarchy is a primary mechanism for establishing relationships. *Can RDF sufficiently express or capture the contingencies and nuances arising from a glocal, pluralized framing?* If nothing else, the test cases that archives are in a position to present to RDF developers may help to complexify an approach that has until recently been premised upon simplicity of expression and flatness of relationships.

Conclusion: some possible approaches

So how should we go about effecting this reorientation of archival conceptualizations and practices? The following are some possible strategies that have been proposed in recent work:

- *Expanding the conception of provenance and associated ideas about authorship, authority, and ownership.* Co-creatorship and the closely related constructs of multiple simultaneous and parallel provenance are propositions that challenge how authorship is traditionally ascribed to archival materials and that seek to complexify how it is represented in provenance-based archival description and archival authority files (Hurley 2005a & b). In both cases, "With their emphasis on a single creating entity, they fail to acknowledge that multiple parties with different types of relationships to each other can be involved in the genesis of records, and be in different types of relationships to each other. They argue, for example, that subjects as well as creators of records should be acknowledged as participants in that genesis, and that archivists have an ethical imperative to pursue descriptive mechanisms for representing both creator and co-creator worldviews and experiences, and supporting diverse user needs and concerns, within and relating to a given community of records" (Gilliland 2013 in press).
- *Focusing on archival information retrieval (IR).* While archival retrieval has not received much attention from IR researchers, there are significant potential advances that have been identified as fertile aspects for research and development. Many of these are focused on XML-based retrieval and they include "helping users to find previously unknown and possibly 'smoking gun'–type documents; establishing the meaningful absence (as opposed to the presence) of documents; and exploiting multiple types and sources of metadata" in order to support different user needs and ontological approaches (Furner and Gilliland 2014 forthcoming).
- *Pluralizing archival education and pedagogy.* The Archival Education and Research Initiative (AERI)'s Pluralizing the Archival Curriculum Group have laid out an archival education framework "to promote a critique of professional and societal norms and include diverse perspectives on archival theory and practice" (PACG 2011). This framework lays out

eight objectives that are accompanied by suggested curricular and pedagogical strategies and cultural competencies designed to prepare archival professionals to work with a pluralized and glocal orientation.

Finally, a new initiative is underway as an outcome of a recent symposium in human rights and archives to develop a framework similar to that of the PACG but addressing the development of metadata infrastructures capable of empowering users seeking to use archives for human rights purposes.

No single strategy, however, will be sufficient to fundamentally change how the archival field approaches the development of an access infrastructure that can meet glocal demands. It will take cooperation across archival institutions worldwide; consultation and equitable collaboration with diverse communities; fundamental shifts in how archival principles such as provenance and hierarchy in description are construed; partnerships with other information and memory institutions, input into the development of the Semantic Web; a strong commitment to research and development, especially in the areas of user community and needs analysis, interface design, and information retrieval; and the fostering of glocal perspectives and skills in all of these areas through archival education.

References

The Aboriginal and Torres Strait Islander Library and Information Resources Network (ATSILIRN). ATSILIRN Protocols. (2012). http://www.aiatsis.gov.au/atsilirn/protocols.php (23 October, 2013)

Archival Education and Research Institute (AERI), Pluralizing the Archival Curriculum Group (PACG). Educating for the Archival Multiverse. *The American Archivist*. (Spring/Summer 2011); 68-102

DigCCurr: Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum. About Our Projects. (n.d.). http://ils.unc.edu/digccurr/ (23 October, 2013)

First Archivist Circle. Protocols for Native American Archival Materials. (2007). http://www2.nau.edu/libnap-p/ (23 October, 2013)

Flinn, Andrew. Archival Activism. Independent and Community-led Archives, Radical Public History and the Heritage Professions. *InterActions: UCLA Journal of Education and Information Studies* 7 (2011), 2. http://www.escholarship.org/uc/item/9pt2490x (23 October, 2013)

Furner, Jonathan; Gilliland, Anne J. Archival IR: Applying and Adapting Information Retrieval Approaches in Archives and Recordkeeping Research, in Research in the Archival Multiverse, Caulfield : Social Informatics Monograph Series, Monash University Press

Gilliland, Anne J. Conceptualizing Twenty-first-century Archives. Chicago : Society of American Archivists, 2013 (in press)

Gilliland, Anne J. Neutrality, Social Justice and the Obligations of Archival Educators and Education in the Twenty-first Century. *Archival Science* 11 (2011), 3-4; 193-209

Gilliland, Anne J.; Lau, Andrew J; McKemmish, Sue. Pluralizing the Archive, chapter 7 in Archives for Maintaining Community and Society in the Digital Age*,* Keiji Fujiyoshi, ed. Kōya : Koyasan University, Japan, 2013; 65-74

Harris, Verne. Archives and Justice: A South African Perspective. Chicago : Society of American Archivists, 2007

Higgins, Sarah. Digital Curation: The Emergence of a New Discipline. *International Journal of Digital Curation*. 2 (2011), 6; 78–88

Hurley, Chris. Parallel Provenance: (1) What, if Anything, is Archival Description? *Archives and Manuscripts* 33 (2005), 1; 110-45

Hurley, Chris. Parallel Provenance: (2) When Something is *not* Related to Everything Else, *Archives and Manuscripts* 33 (2005), 2; 52-91. http://www.infotech.monash.edu.au/research/groups/rcrg/publications/parallel-provenance-combined.pdf (23 October 2013)

Jimerson, Randall C. Archives Power: Memory, Accountability, and Social Justice. Chicago : Society of American Archivists, 2009

Robertson, Roland. Comments on the 'Global Triad' and 'Glocalization.' *Institute for Japanese Culture and Classics, Conference on Globalization and Indigenous Culture*, 1997, http://www2.kokugakuin.ac.jp/ijcc/wp/global/15robertson.html (23 October, 2013)

Robertson, Roland. The Conceptual Promise of Glocalization: Commonality and Diversity. *ART-e-FACT: Strategies of Resistance*. 4 (n.d.). http://artefact.mi2.hr/_a04/lang_en/theory_robertson_en.htm (23 October, 2013)

Society of American Archivists. Core Values Statement. (May 2011). http://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics (23 October, 2013)

Upward, Frank. Structuring the Records Continuum Part One: Postcustodial Principles and Properties. *Archives and Manuscripts* 24 (1996); 268–285

Upward, Frank. Structuring the Records Continuum Part Two: Structuration Theory and Recordkeeping. *Archives and Manuscripts* 25 (1997); 10–35

Upward, Frank; McKemmish, Sue; Reed, Barbara. Archivists and Changing Social and Information Spaces: A Continuum Approach to Recordkeeping and Archiving in Online Cultures. *Archivaria* 72 (Fall 2011); 197-238

Recordkeeping and Archiving in the Cloud. Is There a Silver Lining?

Sue McKemmish
Centre for Organisational and Social Informatics
Monash University, Australia
sue.mckemmish@monash.edu

Summary

There is a rapid uptake of cloud computing in many places around the world. What are the implications for recordkeeping and archiving? Cloud computing offers attractive benefits including significant cost savings, efficiencies, flexibility and scalability, as well as opportunities for the innovative development and delivery of new records management and archival services. Recordkeeping and archiving in the cloud also carry significant risks associated with security, privacy, integrity, authenticity, accessibility and digital continuity, as well as issues relating to commercial continuity and the lack of transparency of cloud services. This paper provides an overview of the current cloud computing environment, different models and types of cloud services, and related recordkeeping and archiving benefits and risks. It uses a case study approach to explore the strategies that two archival authorities in Australia are pursuing in their role as records management standard setters. They include risk assessment approaches and the development of checklists and other tools to guide the evaluation and selection of cloud services, risk management, and contract negotiation. The paper also briefly references the European Union's Cloud for Europe initiative. Finally it calls on recordkeeping and archiving communities to take a pro-active approach, as both consumers and potential service providers, to influencing the future of recordkeeping and archiving in the cloud.

Key words: cloud computing, recordkeeping, archiving, risk management

Introduction

This paper explores the question: what are the implications for recordkeeping and archiving of the rapid uptake of cloud computing? It provides an overview of the current cloud computing environment, different models and types of cloud services, and related recordkeeping and archiving benefits and risks. It uses a case study approach to explore the cloud computing strategies that the National Archives of Australia and the Public Record Office of Victoria are pursuing in their role as records management standard setters. The case study is based on analysis of the policies, strategies and guidelines for government organizations developed by archival authorities to optimize the benefits and miti-

gate the risks associated with records in the cloud. National and state government archival authorities in Australia have a dual role in relation to the regulation of current recordkeeping for good governance and democratic accountability, as well as preservation of government archives as part of Australia’s collective memory and cultural heritage. The case study focuses on their role in relation to the regulation and promotion of current recordkeeping. However, the two roles are inter-related as best practice recordkeeping, particularly in today’s digital environments, is critical to the long-term preservation of archival records of continuing value. Over the past few years the National Archives of Australia and Public Record Office of Victoria have developed strategies relating to cloud computing which are based on risk assessment and management approaches. The paper goes on to present a categorization of risks, and a list of questions relating to cloud services for organizations planning to put their records in the cloud. It is based on an analysis of the policies, guidelines, checklists and other tools issued by these two archival authorities to guide the evaluation and selection of cloud services, risk assessment and management, contract negotiation and the monitoring of cloud services, Finally the paper references recent developments in the European Union, in particular the Cloud for Europe initiative. In conclusion, it calls on recordkeeping and archiving communities to take a proactive approach, as both consumers and potential service providers, to influence the future of recordkeeping and archiving in the cloud.

Records are defined in this paper in accordance with AS/NZ ISO 30300 as:

> Information created, received and maintained as evidence and as an asset by an organization or a person, in pursuit of legal obligations or in the transaction of business (ISO 2012, p. 9).

Archives are defined as records of continuing value. Digital records and archives take multiple forms, come in many formats and are captured in many places, including email systems, websites, blogs, wikis and other social media, business systems including databases, and multimedia systems, as well as Electronic Document and Records Management Systems. They are stored in personal and corporate devices and equipment (on desktops, laptops, tablets and mobile phones, organizational servers), in data stores and digital repositories, including archival digital repositories. In organizational settings, their storage and management might be in-house or outsourced to external service providers. Increasingly records may be managed and stored in the cloud.

Cloud Computing

The US National Institute of Standards and Technology (NIST) defines cloud computing as:

> a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provi-

> sioned and released with minimal management effort or service provider interaction (Mell & Grance 2012, p. 2).

Cloud computing offers attractive benefits including significant cost savings, efficiencies, flexibility and scalability, as well as opportunities for the innovative development and delivery of new services. It also carries significant risks associated with the security, privacy, integrity, authenticity, accessibility and digital continuity of data and records in the cloud. There are also issues relating to commercial continuity and the lack of transparency of cloud services that impact on recordkeeping and archiving.

Public cloud computing offers a pay-as-you-go business model as an attractive alternative to large-scale capital investment in software, platforms and infrastructure, and paying an in-house workforce to manage them. Software-as-a-service (SaaS) delivers business applications hosted by a provider over the web. Platform-as-a-Service (PaaS) provides custom application development or deployment environments in which applications can be built and run on service provider systems. Infrastructure-as-a-Service (IaaS) provides virtual infrastructure components such as servers, storage and network access. Computing resources are accessible everywhere, anytime through diverse media and devices and thus support a mobile and flexible workforce. A cloud services provider pools and dynamically configures its computing resources to meet the needs of multiple clients, enabling rapid scaling of service to meet demand and optimised use of resources.

The rapid uptake of cloud computing is evidenced by the widespread personal and business use of public cloud services such as Microsoft drop boxes, Google drive, Apple's iCloud, gmail, Facebook and other social media, as well as contracted public, private and community cloud services customized to individual business needs. Public cloud services operate on a "multiple tenant" model over the public internet with facilities shared by multiple users. In a private cloud an organization has exclusive use of cloud infrastructure, and in a community cloud, several organizations from a community with common concerns (e.g. related to security, privacy, accountability or jurisdiction) share cloud infrastructure (Mell & Grance 2012, p. 3). Private and community clouds may be managed internally or by a third-party and hosted internally or externally. If they are managed and hosted externally by a third party, private and community cloud services may share more of the benefits and be vulnerable to many of the risks associated with public cloud computing. A hybrid cloud is composed of two or more clouds (public, private or community) and may deliver the best of both worlds. Hybrid clouds and to a lesser extent community clouds potentially enable organisations to realise the benefits of a public cloud (multi-tenancy and pay-as-you-go), but with the added level of privacy, security, accountability and standards compliance usually associated with a private cloud.

Risks associated with the processing, storage and management of data, and more specifically records in the cloud, particularly the public cloud, include security, privacy, integrity, authenticity, accessibility and digital continuity, as well as issues relating to commercial continuity and the lack of transparency of cloud services. The degree of risk and possible consequences for data and records vary for different models and types of services, with the highest risks and most serious consequences generally associated with software and platform services in the public cloud, and relatively low risks associated with infrastructure services. However some of the risks associated with the location of data stores and servers, and trans border data flows are the same regardless of the model or type of service.

The market-speak and images associated with cloud computing provide us with a somewhat misleading mental model which belies the physical reality of the vast server farms owned by Google (US, Finland, Russia and Germany, expanding into Chile, Hong Kong, Taiwan and Singapore), Microsoft (US, Hong Kong, Singapore, Ireland, Brazil), Amazon (US, Dublin, Brazil, Japan), Facebook (US, Sweden), and Apple (USA). From another perspective “cloud” may be an apt term – as there are concerns about the extent to which the operations of service providers are “clouded” in secrecy. There has been a lack of transparency and accountability. For example, many large service providers have policies on non-disclosure of the location of data farms (apparently on security grounds, although server farms are hard to hide as apparently the largest ones are visible from space). More specifically exactly where data is stored, and whether it is being moved around and transferred across borders may not be disclosed to you (see for example Microsoft’s explanation at http://www.microsoft.com/online/legal/v2/en-us/MOS_PTC_Geo_Boundaries.htm). The use of the terms public and private cloud computing, especially in marketing, can also be opaque as the model offered may not always be what it purports to be:

> Be sure that the deployment model offered is what it appears to be and not a marketing ploy whereby a vendor offers differently priced packages of the same services. (PROV 2012, p. 16)

Australian Responses

In 2013, a whole of government policy embracing cloud computing was issued by the Australian Government Information Management Office (AGIMO 2013). The policy was framed as part of the National Digital Economy Strategy and National Cloud Computing Strategy, and highlighted the potential contribution of cloud computing to the national economy, and the leadership role the Australian Government should play in the adoption of cloud computing in all sectors. The policy aims to “help agencies adopt cloud computing to boost innovation and productivity”, and places on government agencies “an explicit obligation to consider cloud services when procuring their new ICT requirements”

(p. 3), taking into account best value for money and adequate management of associated risks. It also prescribes the migration of public web sites to public cloud services:

The Australian Government will be a leader in the use of cloud services to achieve greater efficiency, generate greater value from ICT investment, deliver better services and support a more flexible [mobile] workforce. (AGIMO 2013, p. 3)

In terms of future directions, in line with developments in other parts of the world, the policy points to the possible adoption of a more centralised approach in the form of an Australian Government community cloud, and the possible development of whole-of-government arrangements whereby cloud service providers would submit tenders to become preferred suppliers. In the tenders they would need to demonstrate that they could meet Australian Government requirements. In the meantime, individual government agencies are directed to a range of documents that provide advice and recommendations relating to assessing and managing the risks associated with cloud computing, particularly involving the security and privacy of data processed and stored offshore. These documents include the *Protective Security Policy Framework*, the Defence Signals Directorate's paper on *Cloud Security Considerations* (Department of Defence 2011), and AGIMO's guidelines relating to *Privacy and Cloud Computing for Australian Government Agencies*, and *Negotiating the Cloud: Legal Issues in Cloud Computing Agreements* (AGIMO 2013, pp. 11-12). Following these recommendations and guidelines, personal data and security classified data should only be processed, stored and managed by cloud services located in Australia *and* under Australian jurisdiction – otherwise compliance with Australian security requirements and privacy law cannot be guaranteed.

AGIMO's policy also points to the National Archives of Australia's policy on *Records Management in the Cloud* (NAA 2011a) which was developed in the context of an Australasian Digital Recordkeeping Initiative (ADRI) advisory paper on identifying risks associated with cloud computing, assessing the level of risk for different records (for example secret, confidential, commercially sensitive and personal records), how to perform due diligence when selecting a provider, establishing contractual arrangements, and monitoring their implementation (ADRI 2010). The National Archives' policy states that:

> Cloud computing poses both benefits and risks for Australian Government agencies. Gains in cost and efficiency need to be weighed up against the risks associated with privacy, security and records management.

For records management, it is essential for agencies to consider:

- where records will be stored – there may be risks to Australian Government records when they are stored outside Australia

- the value and nature of the records – the higher the value of the records the more control there needs to be over their management to ensure their integrity, authenticity and reliability
- the risks that may arise – different models of cloud service provision will present different risks to records
- whether risks can be satisfactorily mitigated – this may depend on the ability to negotiate contracts and agreements that address the risks and meet legislative obligations (NAA 2011a, np).

The National Archives policy specifies the need for informed risk assessments and NAA provides a *Check List* (NAA 2011b) for use in assessing risks, assessing and selecting service providers and negotiating contracts. It also specifies the obligations of government agencies to ensure that records in the cloud are:

- governed by the Archives Act 1983
- authentic, accurate and trusted
- complete and unaltered
- secure from unauthorised access, alteration and deletion
- findable, readable and returnable
- related to other relevant records (NAA 2011a, np).

State government archival authorities around Australia have also issued policies and/or guidelines relating to cloud computing. The Public Record Office of Victoria is leading the way. On the basis of extensive research on the implications of cloud computing for recordkeeping (PROV 2012), the Office has issued an excellent *Cloud Computing Guideline* (PROV2013a) and a set of very useful *Tools*, including a risk assessment template and matrix for cloud computing environments, a contract checklist, a requirements checklist and a mapping of intersecting requirements which affect decision making (PROV 2013b). The *Guideline* states that the appropriate treatment of records in the cloud should shape decisions relating to:

- The level of service required
- The contract conditions imposed
- The audit regime adopted
- Restrictions on the physical location of the cloud servers
- Restrictions on the selection of providers by country of registration
- Restrictions on the kind of cloud environment selected (public, private or community) (PROV 2013a, p. 5).

In developing strategies, policies and guidelines for the cloud, the archival authorities frequently reference the suite of international, national and archival authority standards, policies, strategies and guidelines that are already in place relating to digital records and recordkeeping. There is a hidden assumption that best practice digital recordkeeping is already implemented in government organizations, and that good quality records and metadata are already being cre-

ated and managed in digital systems – that the organization's recordkeeping and records are "cloud ready". However as evidenced in reports by Audit Offices, Ombudsman's Offices, Privacy Commissioners and other watchdogs in a number of Australian jurisdictions, there are many cases of poor quality recordkeeping in digital environments, in particular a lack of recordkeeping functionality in business systems and databases.

The Recordkeeping Implications

A close analysis of the issues papers, policies and guidelines published by the National Archives and Public Record Office of Victoria highlights a range of risks relating to processing, storing and managing records in the cloud (ADRI 2010, NAA 2011, 2013a, 2013b, PROV 2012, 2013a, 2013b). On the basis of the analysis, seven risk categories have been identified, and two checklists of questions have been developed to assist government organizations planning to put their records in the cloud, particularly the public cloud. They are presented in Table 1 below. One checklist of questions is designed to help an organisation to evaluate whether potential service providers can meet an organization's requirements relating to their records in the cloud. The other checklist of questions relate to the "cloud readiness" of an organization's recordkeeping management frameworks and digital recordkeeping.

Table 1. Recordkeeping Risks and Checklists

RISK CATEGORIES	CHECKLIST FOR SERVICE PROVIDER	CHECKLIST FOR ORGANIZATON
Location and Legal Jurisdiction Note: Risks in this category are particularly relevant to personal, security and confidential records	Can the service provider disclose precise information about location and movement of data/records? Can the service provider negotiate a contract that guarantees compliance with the laws in the client's jurisdiction so that the client can meet legal obligations re privacy, security, FOI and legal disclosure?	Does the organization know what laws apply to data, information and records in the jurisdiction(s) in which its records will stored (e.g. evidence laws, disclosure laws, privacy laws, FOI laws, national security laws)? Would those laws apply to the organization's records? Do the organization know whether those laws put it in breach of the laws in its own jurisdiction (e.g. privacy, disclosure, FOI, data security)? Does the organization know, for example, that the US Patriot Act applies to all data in storage facilities and server farms operated by US companies, regardless of where in the world they are located? Can the organization identify high risk records in terms of security,

		privacy and confidentiality? Does the organization have policies and risk management strategies relating to its use of social media in public clouds?
Transparency **Accountability** **Governance** Note: relevant to all records.	Are the operations of the service provider transparent? What governance, accountability, internal and external auditing and reporting arrangements are in place? Are third party sub-contractual agreements compliant with client requirements? Are there risk management strategies in place to mitigate unauthorised actions on records, security breaches, disasters resulting in loss or damage, the possibility of cyber criminals, terrorists, or spies hacking or scraping service provider systems?	Does the organization have in place a policy and processes for contract negotiation and monitoring compliance with contractual arrangements?
Protection of Rights in Records Note: This category is particularly relevant to personal and confidential records, and records that might be subject to copyright and intellectual property rights.	Who owns records in the cloud service? Is the privacy of data subjects protected as required by the law in the client’s country, including trans-border movement of data? How is third party access to client records managed, for example if required by a government watchdog organization in the jurisdiction in which the records are stored? Can the service provider support disclosure and access according to the rights regime in the client agency’s jurisdiction?	Can the organization specify its rights management requirements relating to records? Can the organization identify which records might include information that requires rights management?
Recordkeeping Functional Requirements Note: Typically, requirements specify that records are authentic, accurate, reliable, complete, discoverable, accessible, retrievable, readable, and persistently linked to metadata relating to their content, structure, context, management and use.	Do service provider systems have recordkeeping functionality to manage records and their metadata? Do systems create and manage metadata about service provider action on records to ensure their authenticity and integrity? Can systems ingest, migrate, convert, refresh, destroy and extract records according to client specified standards? Do service provider systems protect records from unauthorised	Can the organization specify recordkeeping functionality and metadata requirements for the different types of records to be moved to the cloud? Do in-house systems meet recordkeeping functional requirements? Are records in in-house systems and applications “cloud ready”? Are organizational systems capable of exporting and ingesting records with their metadata into/from storage and management systems in the cloud?

This category is relevant to all records and their metadata.	actions, e.g. illegal destruction, unauthorised access and use, hacking or scraping, and security breaches? Do they have best practice disaster recovery functionality? Can the service provider guarantee that no copies of client records are retained after termination of the contract? Can the service provider return all records and associated metadata in specified formats?	
Digital Continuity Note: particularly relevant to records of continuing value.	Can service provider preserve authentic records over long periods of time?	Can the organization specify its archival preservation requirements?
Vendor Lock-in Note: relevant to all records.	Does the service provider use proprietary systems and formats that potentially lock data into the service provider's servers?	Can the organization specify its requirements relating to proprietary/open source software and formats?
Commercial Continuity Note: relevant to all records.	What contingency plans does the service provider have in the event that it goes out of business or is taken over by another company?	Does the organization have adequate due diligence strategies and procedures to check the commercial viability of service providers and contingency planning?

Cloud for Europe

There has been growing concern in the European Union with the consequences of the fragmentation of the European public sector cloud computing market – namely that EU requirements relating to data ownership, integrity, preservation, data protection, privacy, transparency and accountability have little impact, services are not well integrated, and public sector clients do not get the best value for money.

> [A EU Justice] Opinion examines issues associated with the sharing of resources with other parties, the lack of transparency of an outsourcing chain consisting of multiple processors and subcontractors, the unavailability of a common global data portability framework and uncertainty with regard to the admissibility of the transfer of personal data to cloud providers established outside of the EEA. Similarly, a lack of transparency in terms of the information a controller is able to provide to a data subject on how their personal data is processed is highlighted in the opinion as matter of serious concern (EU Justice 2012, p. 2)

In response, as part of the Digital Agenda for Europe, the European Cloud Partnership (ECP) of public sector users and industry experts aims to shape the

market in Europe and stimulate a European cloud industry that can meet public sector requirements.

> The ECP aims at driving the first steps towards better public procurement of cloud services in Europe, based on common definitions of requirements and possibly eventually going as far as joint procurement across borders …
>
> Pooling public requirements could bring higher efficiency and common sectoral requirements (e.g. eHealth, social care, assisted living, eGovernment services) would reduce costs and enable interoperability. The private sector would also benefit from higher quality services, more competition, rapid standardisation and better interoperability and market opportunities for high-tech SMEs (European Commission 2013).

The Partnership has put in place a range of consultation strategies to engage stakeholders in determining the requirements of the Cloud for Europe, presenting an opportunity for the pro-active engagement of archival institutions, recordkeeping standards setters and the recordkeeping and archiving community in ensuring that requirements relating to records and archives in the cloud are fully addressed.

Towards Archives 3.0

Stančić, Rajh and Milošević (2012) introduce the concept of Archiving-as-a-Service and the need to transition beyond "custody" from a "postcustody" to a "postcustody 2.0" paradigm as illustrated in Figure 1 (which reproduces Figure 16 (p. 124) with the kind permission of the authors).

They put forward four scenarios for archiving in the cloud:

1. Service providers are responsible for control of archived content without much interference and additional control taken by creators and archival institutions;
2. Creators invest much effort in additional control of non-standardized services;
3. Services are standardized through best practices and creators recognize the importance of choosing providers consistent to these practices;
4. Archival community is actively involved in the new concept of archiving and influence providers' practices.

They conclude that the fourth "postcustody 2.0" scenario is "probably the best way to ensure long-term protection, preservation and usage of electronic content created and archived today" (p. 123), and urge the archival community to be proactive "in the formation of the new, preservation-aware cloud services" (p. 124). The Australian case study discussed earlier in this paper is an example of a "postcustody 2.0" approach. In the Australian context, government archival institutions are engaged in regulating and advising government agencies. Through the standards they set for records in the cloud, they also hope to influence the development of whole-of-government preferred supplier procurement

arrangements, and the cloud services offered by providers. At this stage in the evolution of cloud services, no Australian government archival institutions would sanction the placement of records identified as being of continuing value in the cloud. However, it is possible to envisage future scenarios where archival records might be managed and stored in the cloud, for example the establishment of an Australian government community cloud in line with developments in Europe, the UK and Canada, in this case serviced by providers compliant with the relevant laws in Australian jurisdictions.

Figure 1. Changes of Archival Practice

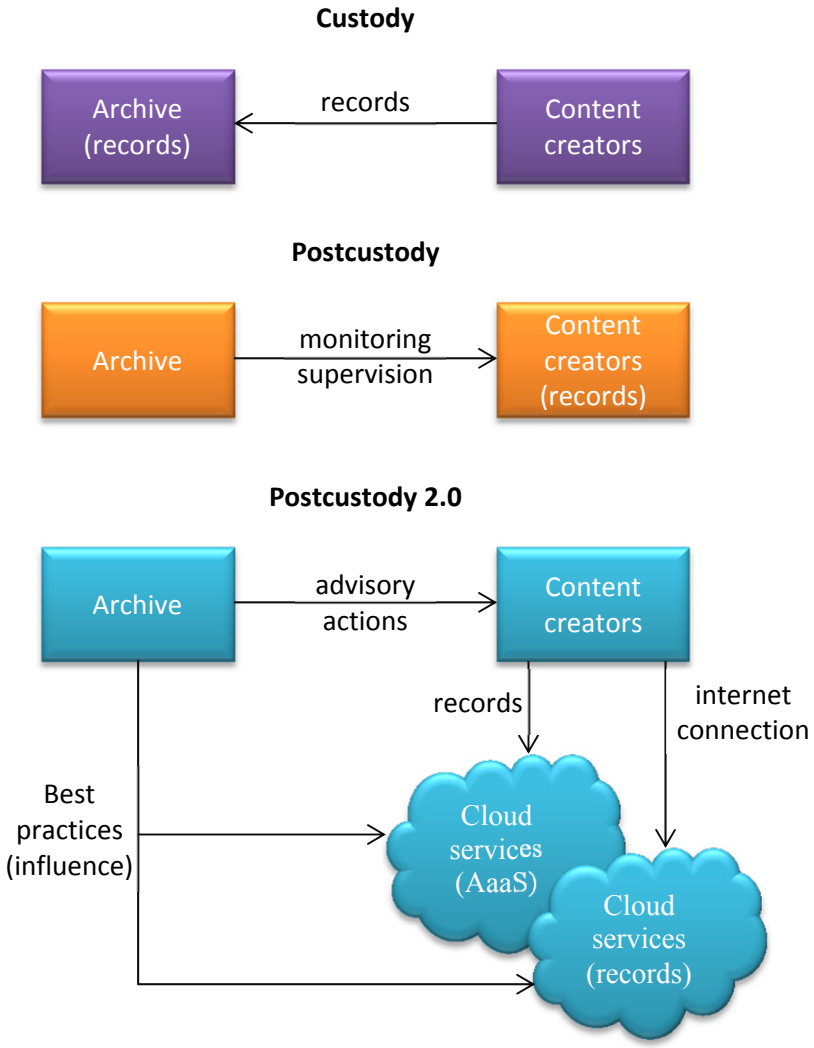


Figure 16. Changes of Archival Practice

Conclusion

Somewhere beyond custody, in the archival multiverse, lies the vision of Archives 3.0, taking advantage of the benefits and opportunities of cloud computing to build community clouds, e.g. of national, state and community archives in the EU or Australia. In this scenario the policies, standards, strategies, guidelines and tools being developed by archival authorities like the National Archives of Australia and Public Record Office of Victoria contribute to the development of broader requirements that embrace the recordkeeping and archiving needs of all those involved in the community cloud partnership. To realize this vision, recordkeeping and archiving communities need to take a pro-active approach, as standards setters, clients, and potential service providers, to shaping cloud services, and the future of recordkeeping and archiving in the cloud.

References

Australian Department of Defence. Cloud Computing Security Considerations. Canberra: Australian Government, 2011.

Australasian Digital Recordkeeping Initiative. Advice on Managing the Recordkeeping Risks Associated with Cloud Computing. Canberra: ADRI, 2010. http://www.adri.gov.au/products.aspx (12 October 2013).

Australian Government Information Management Office (AGIMO). Australian Government Cloud Computing Policy: Maximising the Value of the Cloud v. 2.0. 2013. http://agict.gov.au/sites/default/files/Australian%20Government%20Cloud%20Computing%20Policy%20Version%202.1.pdf (9 October 2013)

European Commission. Justice Data Protection WP Article 29. EU Justice Opinion 05/2012 on Cloud Computing. 2012. http://ec.europa.eu/justice/data-protection/article-29/documentation/index_en.htm (12 October 2013)

European Commission. Digital Agenda for Europe, European Cloud Computing Strategy – CloudforEurope. 2011. http://www.cloudforeurope.eu (12 October 2013)

Mell, P; Grance, T. The NIST Definition of Cloud Computing, National Institute of Standards and Technology. 2010. http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf (1 October 2013)

National Archives of Australia. Records Management and the Cloud. NAA, Canberra: NAA, 2011a. http://www.naa.gov.au/records-management/agency/secure-and-store/rm-and-the-cloud/ (12 October 2013).

National Archives of Australia. A Checklist for Records Management and the Cloud. Canberra: NAA, 2011b. http://www.naa.gov.au/records-management/publications/cloud-checklist.aspx (12 October 2013).

Public Record Office Victoria (PROV). PROV Cloud Computing Guideline 1: Cloud Computing Decision-Making. North Melbourne: PROV, 2013a. http://prov.vic.gov.au/wp-content/uploads/2013/06/Cloud_Computing_Guideline_1.pdf (12 October 2013).

Public Record Office Victoria (PROV) 2013, PROV Cloud Computing Guideline 2: Cloud Computing Tools. North Melbourne: PROV, 2013b. http://prov.vic.gov.au/wp-content/uploads/2013/06/Cloud_Computing_Guideline_2.pdf (12 October 2-13)

Public Record Office Victoria (PROV). 2013, PROV Cloud Computing Policy. North Melbourne: PROV, 2013c. http://prov.vic.gov.au/wp-content/uploads/2013/06/Cloud_Computing_Policy.pdf (12 October 2013)

Public Record Office of Victoria (PROV). Cloud Computing Implications for Records Management. North Melbourne: PROV, 2012. http://prov.vic.gov.au/wp-content/uploads/ 2012/04/Issues-Paper-Cloud-Computing.pdf (12 October 2013)

Stančić, Hrvoje; Arian Rajh, Arian; Milošević, Ivor. "Archiving-as-a-Service". Influence of Cloud Computing on the Archival Theory and Practice. // *The Memory of the World in the Digital Age: Digitization and Preservation* / Duranti, Luciana ; Shaffer, Elizabeth (ed). UNESCO, 2013, 108-125.

Standards Australia; Standards New Zealand. AS/NZ ISO 30300: 2013: Information and Documentation – Management Systems for Recordkeeping – Fundamentals and Vocabulary (ISO 30300: 2011 MOD). Sydney: Standard Australia, 2012; Wellington: Standards New Zealand, 2012.

From Translation Machine Theory to Machine Translation Theory – some initial considerations

Oliver Čulo
Universität Mainz
Faculty 06 of Johannes Gutenberg University Mainz
Postfach 11 50, D 76711 Germersheim, GERMANY
culo@uni-mainz.de

Summary

Early, rule-based machine translation worked on the basis of contrastive linguistic and translational considerations. The advent of corpus-based machine translation techniques from the 1980s on shifted this paradigm towards a more computationally oriented one. Lately, through the use of techniques like domain adaption or factored models, linguistic and translational considerations have made a comeback. Still, the exchange between the two disciplines of machine translation and translation studies is rare, though they have already benefitted from each other at least in one way: Translators nowadays regularly use tools like translation memories. Present-day machine translation can benefit from translation theory, too, as this paper attempts to argue, by incorporating theoretical and experimental insights into areas like factors that guide translation shifts. For this, it is necessary for both to share at least some common terminological code and theoretical perspective. This paper will attempt to lay a basis for that by locating relevant concepts from machine translation within translation theory and vice versa, and investigating how the two fields could further benefit from each other.

Key words: translation studies, machine translation, translation theory, translation shifts, translation factors, post-editing

Introduction

Early, rule-based Machine Translation (henceforth MT) worked on the basis of contrastive linguistic and translational considerations. Dictionaries were hand-built and rules were manually created in order to transfer syntactic structures from one language into the other. The basis for transfer often was some kind of valency information or dependency structure, allowing syntactic complements to be shifted around and adapted for differences e.g. in word order or morphology between languages. Examples of this are for instance EUROTRA (Johnson, King, and des Tombe 1985; Copeland et al. 1991) or the METAL system (Gebruers 1988); Hutchins and Somers present an overview of early MT systems (Hutchins and Somers 1992). Present-day MT, however, is mainly based

on a computationally oriented, statistical model of translation (Brown et al. 1993), with techniques like sentence alignment (Gale and Church 1993), word alignment (Och and Ney 2003) or phrase alignment (Koehn, Och, and Marcu 2003) – where phrases are not phrases in the strictly linguistic sense, but continuous stretches of words – as key techniques that research is performed on. However, as the following sections will show, MT has lately begun to re-integrate linguistic as well as translational knowledge. Also, a new field is emerging in which human translators and machine translators meet: post-editing. Given these recent developments, this paper argues for an intensified exchange between the two fields of translations studies (TS) and MT. In the following, I present some initial considerations how MT could benefit from TS research done on text typology and post editing. However, before such an attempt is made, it is necessary to locate MT within the field of translation research, both to understand what the relation of MT to other TS fields is, as well as to create a basis for a common vocabulary for the two disciplines, which should facilitate exchange.

What kind of translation is MT?

Functional translation theory, in terms of (Nord 1997; Nord 2006), distinguishes two types of translations: documentary and instrumental translations.[1] The documentary translation aims to reflect key features of the original, e.g. choice of words, syntactic structure, by mimicking them in the target language, leaving them more or less unchanged. This could, for instance, be a gloss as it is often given in linguistic literature for non-English examples. The instrumental translation, on the other hand, aims to function as a text that could have been produced as such in the target language as an original, too. This would e.g. be a translated manual or a newspaper report. Instrumental translation involves cases in which the original function of a text is kept (i.e. *functionally equivalent* translations) as well as those cases in which it is changed (*functionally divergent*). Translations can be functionally divergent on purpose or simply because they cannot function in the same way in the target language. For instance, a report on gay rights, which may serve a simple information purpose in the source language and culture, may have a very different effect in a target culture.

We can assume that the main goal of MT is to produce a functionally equivalent translation, while we cannot rule out that functionally divergent translations may be produced using MT, as the product may receive a different functional interpretation in the target language. As for the status in terms of documentary vs. instrumental, high quality MT would certainly be aimed at creating fully instrumental translations. But the current performance of MT systems only allows for the product as being defined as documentary translations. This is, however,

[1] There are, of course, competing distinctions like House's distinction of overt and covert translation (House 1997), but they are comparable in their key characteristics (Nord 2006).

only valid at the textual level. While we can hope to at least get the gist of a machine-translated text, there may still be enough instances, single words, phrases or sentences, in there that are completely incomprehensible.
This being a tentative initial analysis of the status of MT from the viewpoint of TS, it still is a first step to entrench aspects of MT in translation studies such that a common vocabulary and a common point of view is established – two important factors for intensified exchange. That this is not an easy task and will need further work in order to overcome philosophical problems as well as resistance from within the TS community, has been argued e.g. by Rozmyslowicz (in press). The author argues that the fact that MT has not yet been identified as a TS problem, "is not accidental but a symptom of certain theoretical and methodological predispositions" (ibid.). On the other hand, as has been said before, MT of the past decades has been more concerned with solving computational problems and, in my view, needs to become more open towards newer developments in TS. The following sections will briefly assess from what kinds of Ts concepts and research MT could benefit.

Linguistic and translational knowledge in MT

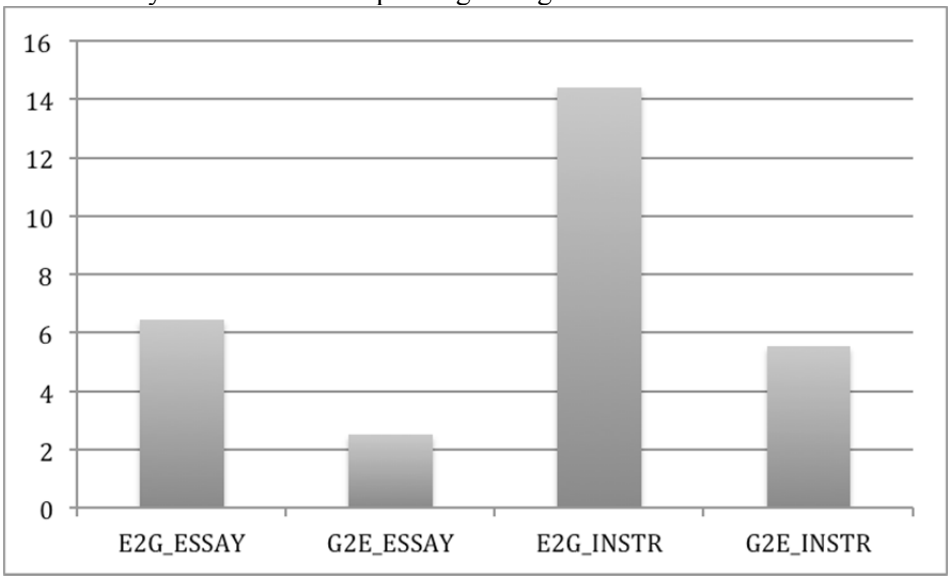
Recently, statistical MT has in many ways re-integrated linguistic knowledge, e.g. by using dependency treelets for transfer (Ding and Palmer 2005; Quirk, Menezes, and Cherry 2005), using factored translation models incorporating morphology, part-of-speech, etc. (Koehn and Hoang 2007), adding semantic roles (Wu and Fung 2009; Haugereid and Bond 2012), or creating hybrid systems using both rule-based and statistical knowledge (Žabokrtský, Ptáček, and Pajas 2008). In all these fields, not only linguistic, but also translation research can contribute to relevant knowledge of the field, for instance in the research on domains and their textual properties, as will be demonstrated later on.
While these systems are using additional linguistic knowledge, few attempts have been made at including truly translational knowledge. Though training models based on domains of discourse (Koehn and Schroeder 2007; Bertoldi and Federico 2009) is not explicitly translational, it does match the affinity of many translation scholars to have a functional look at the text and what context it is embedded in. Domain here is not the only, but a very important factor, where text type is another (Reiß 1976). The notion of *register* from systemic functional linguistics (Halliday and Hasan 1989) can be seen as a crossing between domain and text type and has been made fruitful for translation analysis (Teich 2003; Neumann 2008; Hansen-Schirra, Neumann, and Steiner 2012). The definition of a register takes into account both the field of discourse (i.e. the topic or domain) as well as the mode (e.g. written, spoken); a third variable is the tenor, the relation between those involved in discourse. The type of register a text belongs to has a strong influence on its linguistic characteristics (Biber 1995; Neumann 2008). In combination with translation direction, it is a decisive factor in guiding translation shifts. Figure 1 shows the proportions of shifts be-

tween the grammatical functions subject and adverb of any kind per combination of translation direction (E2G = English to German, G2E = German to English) and register (ESSAY = political essay, INSTR = computer manuals). The bars read as follows: When, for G2E_INSTR, a subject is shifted, in more than 4% of the cases it is shifted to some sort of adverb. We can see from this graph that this proportion is about three times higher in the opposite translation direction which in many cases is due to a typological contrast between English and German. English can fill the subject with inanimate entities which German cannot do easily, often resulting in translation pairs such as the following:

Tray 1 [...] holds up to 125 sheets [...]
In Fach 1 [...] können bis zu 125 Blatt Papier eingelegt werden [...]
`into tray 1 [...] can up to 125 sheets paper inserted be [...]`
(E2G_INSTR_001)

The English inanimate subject *Tray 1* is rendered as locative adverb *In Fach 1* `into tray 1` in German. Such shifts are quite common and, as figure 1 shows, are more typical for the register INSTR in the translation direction English to German, which, from heuristic examination, can partially be attributed to the use of inanimate subjects as in the above example.

Figure 1: Proportions for subject shifts, indicating how often a subject was shifted to any kind of adverb depending on register and translation direction



The conclusions one might draw for MT are straightforward: Training language models according to translation direction *and* register should be beneficial for the performance of the systems, as shift types are heavily dependent on these two factors. A study in this direction was performed by Kurokawa et al. (2009).

Based on the English-French Hansard corpus, Kurokawa et al. compiled a training corpus in which originals and translations were clearly marked. Models were trained according to the translation direction. The authors report that they were able to reproduce equivalent BLEU scores using only a fifth of the training data when these were sorted according to translation direction, as opposed to training on models in which translation direction is being disregarded of. If we take a closer look, these training data are also from a single register, as the corpus used only contains parliament speeches (written-to-be-spoken text type, political domain). In order to determine the combined effect of translation direction and in-register training of MT models, a broader study setup comparing various registers and translation directions would be necessary.

Post-editing: Where humans and machines meet

In addition to studying the products of translation, both MT scholars and translation scholars are becoming more and more interested in the process of post-editing (O'Brien 2002; Tatsumi 2009; Groves and Schmidtke 2009; O'Brien 2010; Carl et al. 2011), i.e. the process of humans correcting MT output. Post-editing (PE) is a field in which the human translator and the machine meet – as well as the two disciplines MT and TS. It is the task of editing erroneous MT output with the goal of getting it either into a just about usable shape (often referred to as *light* or *fast* PE) or making it indistinguishable from proper translations (*full* PE). This task poses specific problems as compared to purely human translations, as the post editors have to deal with output that can be erroneous on multiple levels: morphology, syntax, semantics and last but not least pragmatics. Also, the cognitive load is heightened with respect to focus: When post-editing, translators have to focus on both the source text as well as the MT output.

These challenges are an opportunity at the same time: As many have noted, a monitored PE process can reveal much about the errors an MT system makes and can thus help improve the system. Also, contrasting PE and the "usual", purely human translation helps in understanding more about either kind of process.

Research on post-editing can surely be said to be still in its infancy. So far, research on post-editing focuses on such questions as the different post-editing strategies (Carl et al. 2011), on studying efficiency both with respect to time as well as quality (e.g. Groves and Schmidtke 2009; Specia 2011), on differences between post-edited and from-scratch translated texts (Čulo et al. in print), or on the question which additional skills are needed for post-editing (O'Brien 2002).

PE is subject to a number of factors, e.g. to the familiarity of the translator with the mechanisms of MT, to their technical skills, and to their attitude towards MT. Besides these, the task description of PE is crucial: When do I discard a sentence and start from scratch, when do I use parts of it and correct it, and when is it good enough for my task description? One of the biggest problems is

the question what exactly light PE is: How do I decide whether a sentence is good enough to be understood, which errors can I disregard of etc. A major question, again, is: Which concepts from TS can be re-used and how can PE-relevant concepts be entrenched in TS and in translator teaching? For instance, the pilot study by (Čulo et al. in print) reveals that certain PE strategies like non-backtracking behaviour (Carl et al. 2011) can lead to errors in consistency. In the example discussed by Čulo et al., the English word *nurse*, which is ambiguous with regard to gender, appeared both in its German female as well as male form (*Krankenschwester* vs. *Krankenpfleger*) in the post-edited texts, where the correct interpretation would have been the male form; this mistake did not happen in from-scratch translations by humans. More careful task descriptions with regard to which strategies a translator is familiar with should be pursued might help avoid such mistakes. This, again, being a tentative point of view given here, needs further study. Various ongoing post-editing research projects like one at the Faculty for Translation Studies, Linguistics and Cultural Studies at Mainz University are looking at such factors impacting on PE.

Conclusion: The time is right!

This paper has presented some, admittedly, quite preliminary considerations on where MT and TS can meet, both with respect to theoretical and practical issues, and how both could profit from an intensified exchange.

On a theoretical level, MT can be described in terms of functional theories on translation, and thus formalised as a type of translation which is very typical for a number of settings in which translators work – the goal of a functionally equivalent translation in mind. Given such a theoretical entrenchment of MT in TS, acceptance for MT could be raised within the translation community, also opening various collaboration opportunities in terms of research. On a practical level, MT could benefit from various studies performed in empirical TS on how factors like register or translation direction impact on linguistic properties and translation shifts. Using this knowledge, MT can optimise its strategies in training MT models, reducing the necessary size of data and in consequence the prerequisites for processing power. In terms of the emerging field of PE, the definition of the PE task based on translational concepts can help optimise the task setting, leading to even larger efficiency and quality gains than those already reported.

Future work for this line of research on how MT and TS can further converge will consist of a broader overview and better formalised model on what kind of benefits can be expected and what the preconditions are for each case. Given the current developments in which exchange has already picked up (see references cited in this paper), the time for this has probably not been better for this endeavour than in quite a while.

References

Bertoldi, Nicola, and Marcello Federico. 2009. “Domain Adaption for Statistical Machine Translation with Monolingual Resources.” In *Proceedings of the Fourth Workshop on Statistical Machine Translation*, 182–189. Athens, Greece: Association for Computational Linguistics.

Biber, Douglas. 1995. *Dimensions of Register Variation*. Cambridge University Press.

Brown, Peter E., Stephen A. Della Pietra, Vincent J. Della Pietra, and Robert L. Mercer. 1993. “The Mathematics of Statistical Machine Translation: Parameter Estimation.” *Computational Linguistics* 2 (19): 263–311.

Carl, Michael, Barbara Dragsted, Jakob Elming, Daniel Hardt, and Arnt Lykke Jakobsen. 2011. “The Process of Post-editing: a Pilot Study.” In *Proceedings of the 8th International NLPSC Workshop. Special Theme: Human-machine Interaction in Translation*, edited by Bernadette Sharp, Michael Zock, Michael Carl, and Arnt Lykke Jakobsen, 131–142. Copenhagen Studies in Language 41. Frederiksberg: Samfundslitteratur.

Copeland, Charles, Jacques Durand, Steven Krauwer, and Bente Maegaard, ed. 1991. *The Eurotra Linguistic Specifications*. Studies in Machine Translation and Natural Language Processing 1. Brussels.

Čulo, Oliver, Silke Gutermuth, Silvia Hansen-Schirra, and Jean Nitzke. in print. “The Influence of Post-editing on Translation Strategies.” In *Expertise in Post-editing*.

Ding, Yuan, and Martha Palmer. 2005. “Machine Translation Using Probabilistic Synchronous Dependency Insertion Grammars.” In *Proceedings of the 43rd Annual Meeting of the ACL*, edited by Ann Arbor, 541–8.

Gale, William A, and Kenneth W Church. 1993. “A Program for Aligning Sentences in Bilingual Corpora.” *Computational Linguistics* 19 (1): 75–102.

Gebruers, Rudi. 1988. “Valency and MT: Recent Developments in the METAL System.” In *Proceedings of the Second Conference on Applied Natural Language Processing*, 168–175.

Groves, Declan, and Dags Schmidtke. 2009. “Identification and Analysis of Post-editing Patterns for MT.” In *MT Summit XII: Proceedings of the Twelfth Machine Translation Summit*, 429–436. Ottawa, Canada.

Halliday, M. A.K, and Ruqaiya Hasan. 1989. *Language, Context, and Text: Aspects of Language in a Social-semiotic Perspective*. Oxford: Oxford University Press.

Hansen-Schirra, Silvia, Stella Neumann, and Erich Steiner. 2012. *Cross-linguistic Corpora for the Study of Translations. Insights from the Language Pair English-German*. Berlin: De Gruyter.

Haugereid, Petter, and Francis Bond. 2012. “Extracting Semantic Transfer Rules from Parallel Corpora with SMT Phrase Aligners.” In *Proceedings of the Sixth Workshop on Syntax, Semantics and Structure in Statistical Translation*, 67–75. Jeju, Republic of Korea: Association for Computational Linguistics. http://www.aclweb.org/anthology/W12-4208.

House, Juliane. 1997. *Translation Quality Assessment. A Model Revisited*. Tübingen: Gunter Narr Verlag.

Hutchins, John, and Harold Somers. 1992. *An Introduction to Machine Translation*. London u.a.: Acadamic Press.

Johnson, Rod, Maghi King, and Louis des Tombe. 1985. “EUROTRA: a Multilingual System Under Development.” *Computational Linguistics* 11 (2-3): 155–169.

Koehn, Philipp, and Hieu Hoang. 2007. “Factored Translation Models.” In *Proceedings of EMNLP-CoNLL*, 868–876.

Koehn, Philipp, Franz Josef Och, and Daniel Marcu. 2003. “Statistical Phrase-Based Translation.” In *Proceedings of HLT-NAACL 2003*, 127–133.

Koehn, Philipp, and Josh Schroeder. 2007. “Experiments in Domain Adaptation for Statistical Machine Translation.” In *ACL Workshop on Machine Translation 2007*.

Kurokawa, David, Cyril Goutte, and Pierre Isabelle. 2009. “Automatic Detection of Translated Text and Its Impact on Machine Translation.” *Proceedings. MT Summit XII, The Twelfth Ma-*

chine Translation Summit International Association for Machine Translation Hosted by the Association for Machine Translation in the Americas.

Neumann, Stella. 2008. *Contrastive Register Variation. A Quantitative Approach to the Comparison of English and German*. Saarbrücken: Universität des Saarlandes.

Nord, Christiane. 1997. *Translating as a Purposeful Activity: Functionalist Approaches Explained*. St. Jerome.

———. 2006. “Translating for Communicative Purposes Across Culture Boundaries.” *Journal of Translation Studies* 9 (1): 43–60.

O’Brien, Sharon. 2002. “Teaching Post-editing: a Proposal for Course Content.” In *Sixth EAMT Workshop*, 99–106. Manchester, U.K. http://www.mt-archive.info/EAMT-2002-OBrien.pdf.

———. 2010. “Introduction to Post-Editing: Who, What, How and Where to Next?” In *Proceedings of AMTA 2010*. Denver, Colorado. http://amta2010.amtaweb.org/AMTA/papers /6-01-ObrienPostEdit.pdf.

Och, Franz-Josef, and Hermann Ney. 2003. “A Systematic Comparison of Various Statistical Alignment Models.” *Computational Linguistics* 29 (1): 19–51.

Quirk, Christopher, Arul Menezes, and Colin Cherry. 2005. “Dependency Treelet Translation: Syntactically Informed Phrasal SMT.” In *Proceedings of the 43rd Annual Meeting of the ACL*, edited by Ann Arbor, 271–79.

Reiß, Katharina. 1976. *Texttyp Und Übersetzungsmethode. Der Operative Text*. Kronberg/Ts.: Scriptor-Verlag.

Rozmyslowicz, Tomasz. in press. “Machine Translation: A Problem for Translation Theory.” *New Voices in Translation Studies*

Specia, Lucia. 2011. “Exploiting Objective Annotations for Measuring Translation Post-editing Effort.” In *Proceedings of the 15th Conference of the European Association for Machine Translation*, edited by Mikel L. Forcada, Heidi Depraetere, and Vincent Vandeghinste, 73–80. Leuven, Belgium. http://www.mt-archive.info/EAMT-2011-Specia.pdf.

Tatsumi, Midori. 2009. “Correlation Between Automatic Evaluation Metric Scores, Post-editing Speed, and Some Other Factors.” In *The Twelfth Machine Translation Summit (MT-Summit XII)*, 332–339. Ontario, Ottawa, Canada.

Teich, Elke. 2003. *Cross-linguistic Variation in System and Text. A Methodology for the Investigation of Translations and Comparable Texts*. Vol. 5. Text, Translation, Computational Processing. Berlin/New York: Mouton de Gruyter.

Wu, Dekai, and Pascal Fung. 2009. “Can Semantic Role Labeling Improve SMT?” In , 218–225. Barcelona.

Žabokrtský, Zdeněk, Jan Ptáček, and Petr Pajas. 2008. “TectoMT: Highly Modular MT System with Tectogrammatics Used as Transfer Layer.” In *Proceedings of WMT 2008*.

THEORY AND METHODOLOGY OF INFORMATION GOVERNANCE

The Lifespan of Web references: An Example in Graduate Papers at the Department of Information Sciences in Zagreb

Anina Bauer
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
anina.bauer@gmail.com

Đilda Pečarić
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
dpecaric@ffzg.hr

Summary

An increasing amount of students and scholars use web references as a prime source in their papers. The main concern is that those references have a short lifespan.
In retrospect to that, the aim of this article is to show how many of the web references, gathered in a corpus of 1947 web references within 362 graduate papers at the Department of Information Sciences at the Faculty of Humanities and Social Sciences in Zagreb in the period from 2003 to 2010, are no longer active, i.e. accessible. It was also substantial to know how many of the web references have disappeared in the period between two researches.
The main hypothesis is that there has been an increase in the number of inactive web references within the aforementioned corpus. It is expected that the highest number of inactive web references is present in the graduate papers that have been written during the first half of the analysed period, i.e. more time has passed since those references have been accessed. As well as the availability of the web references, certain bibliographic data were also analysed, where it is predicted that according to the type of web references web pages are more likely to be inactive than any other types such as scientific articles, etc. In response to that, it is expected that the web sites of inactive web references are non-scientific in snature. The analysis of the presence of authors and publishers among inactive web references is also included.

Key words: web reference, lifespan, information sciences, graduate papers

Introduction

An increasing amount of students and scholars use web references as a prime source in their papers, so it has become necessary to analyse them, especially because those references have a short lifespan. The lack of constant availability of those references may be due to infrequent updating of the web sites on which the references are located, or simply because the topic in the reference is out-dated and no longer useable.

One of the reasons why the web references tend to disappear has to do with the absence of a universal definition of web references[1]; disparity in the division of types of web references[2]; inability to, accurately, estimate their lifespan[3]. Even though most web references disappear within two or three years since they have been placed online, not knowing the exact date of their removal from the Internet complicates the calculation of the average lifespan.[4]

In response to the aforementioned findings, a prior research of web references in graduate papers at the Department of Information Sciences at the Faculty of Humanities and Social Sciences in Zagreb (Bauer, Kirin, Turković, 2011) has led to a secondary research regarding the lifespan of the web references used in the initial corpus.

The aim of this article is to point out a trend of 'disappearing' web references, which are used in scientific papers. The results show how many of the web references are no longer active, whether they have been relocated or completely removed from the Internet since accessing the web references during the former research regarding the amount of used web references until the carrying out of the research this article is dealing with.

The main hypothesis is that there has been an increase in the number of inactive web references within the aforementioned corpus since the conduction of the first research, especially because more than a year has passed. By observing the

[1] Halpin, H. (2011), p. 154 and Spinellis, D. (2003), p. 72.: Is it URI (Uniform Resource Identifier) which is defined as a unique identifier of various sources on the Internet, or should one use one of his subsets such as URL (Uniform Resource Locator) for the location or URN (Uniform Resource Name) for the name of the source, or is a completely new system of identifying in order.

[2] Meyer zu Eissen, S.; Stein, B. (2004), p. 2 and Santini, M. (2007), p. 4: Most frequent is the classification according to the type of web references, as well as the origin. Since there is a number of classification schemes which allow a high level of flexibility in defining web references, one can have so called unsorted types of references, which can result in incredible sources.

[3] Research of Ashenfelder, M. (2011) and *What is the Average Lifespan of a Web Site* (2012): Even though most authors agree that the average lifespan of web references is between 44 and 75 days, one most notice that that includes isolated, personal or business web pages, not to mention a big amount of malware pages which drastically reduce it.

[4] *What is the Average Lifespan of a Web Site* (2012).

lifespan of the web references, one can see how fast the information becomes outdated, not to mention whether the information itself alters or if it simply relocates. The research included an analysis during a longer period from 2003 to 2010, where one expects that the highest number of inactive web references can be found in the graduate papers that were written during the first half of the analysed period, i.e. more time passed since those references had been accessed. It was also interesting to see whether the type of the web references, their location and the presence of authors and publishers influence their lifespan. Expectations include that web pages are more likely to be inactive than any other types such as scientific articles or records within electronic databases, that the references which are located in web sites of commercial, non-scientific nature have a shorter lifespan, and that most inactive web references have an author or publisher listed.

Among other things, this research should provide answers in which domain can the web references that disappear most frequently be found, and also according to the criteria that the active web references meet, what can be done to insure their longevity?

Methods

The research was based on the analysis of data that was collected in the previous survey regarding the use of web references in graduate papers at the Department of Information Sciences at the Faculty of Humanities and Social Sciences in Zagreb conducted by the end of June in the year 2011. The corpus included a set of 362 graduate papers, presented in the period between 2003 and 2010, that were available online in the Digital Repository of the Faculty Library, from which 1947 web references were extracted. Since the graduate papers were written by the students from the Department of Information Sciences, disciplines included were: Archive and Documentation Science, Library Science, Information Science and Museology.[5]

While the main object of the preceding study was to find out the amount of web references students use in their graduate papers, the following research formed around discovering which of the used web references were no longer active. The aforementioned study was carried out by the end of January in the year 2013 using the data that had been previously manually entered into a Microsoft Access database containing web references in graduate papers at the Department of Information Sciences at the Faculty of Humanities and Social Sciences in Zagreb.

The availability of the web references was determined according to the URL addresses that were used for the citation of references in the graduate papers. The results show the state of the availability of the used web references, which

[5] see Bauer, A.; Kirin, M.; Turković, M. (2011), p. 70.

was followed by the analysis of results per access year (when the web references were used), as well as the analysis of the type of web references, web sites and the presence of authors and publishers among inactive web references. The general assumption of the availability of the web references and the per year analysis includes the comparison of data recorded in June 2011 and January 2013, while the distinction according to bibliographic data deals with the data recorded in January 2013.

The Lifespan of Web references

For the purpose of this research, 362 graduate papers from the Department of Information Sciences at the Faculty of Humanities and Social Sciences in Zagreb that included 6525 references, from which 1947 (29.8%) were web references, were analysed.[6]

Even though the type of web references was not identified for 36.7% of them, for the purpose of this research, it was not important whether they had adequate bibliographic data or not (further results lay out the assumptions regarding their type), but whether the URL address, which was listed in the papers, was still valid.

In retrospect to that, the main goal was to find out how many of the analysed web references have become inactive once the first research (June 2011) was carried out, and how many more have become inactive until the conduction of the second research (January 2013).

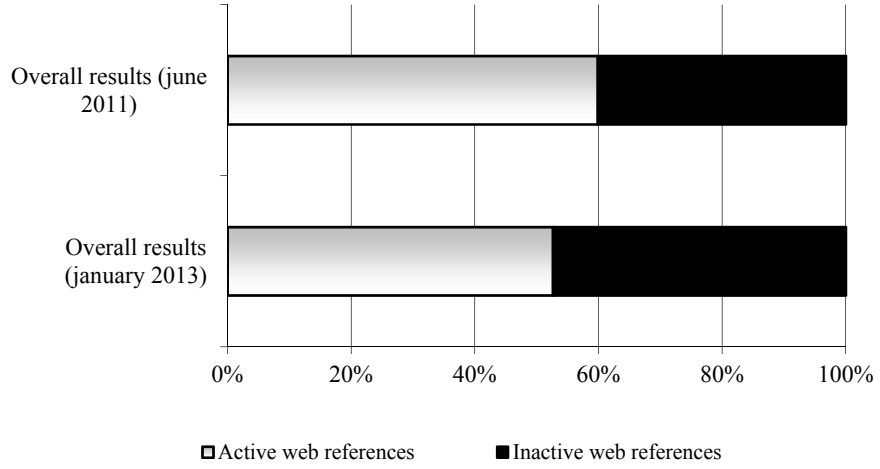


Figure 1. The overall percentage of active and inactive web references

[6] A. Bauer, M. Kirin, M. Turković (2011), p. 70.

As seen in Figure 1, the results show that by the end of June 2011, 781 (40.1%) web references were unavailable, while by the end of January 2013, the number of inactive web references went up to 922 (47.3%).
Considering that the lifespan of web references tends to shorten with time, one expects that there has been an increase in inactive web references by the time of the conduction of this research, and most of them are expected to have disappeared, i.e. been removed from the Internet.
The hypothesis has been confirmed since the number of web references that became inactive in between the two studies has gone up 7.2%, and most of the 922 inactive web references disappeared or became unavailable, while just a few were either relocated, or in some way altered.

Availability of Web references per Year of access
The following included an analysis of the availability of the web references according to the year in which they were accessed. Since not all of the web references have a date of access listed, the year in which the papers, where the references are used, were presented, i.e. published, is used as the access year.
In the first research, among other findings, it was listed how many web references were used in which year. On top of that, for this research it was also included how many of them have been inactive by the end of June 2011 and by the end of January 2013, per year of the analysed period.

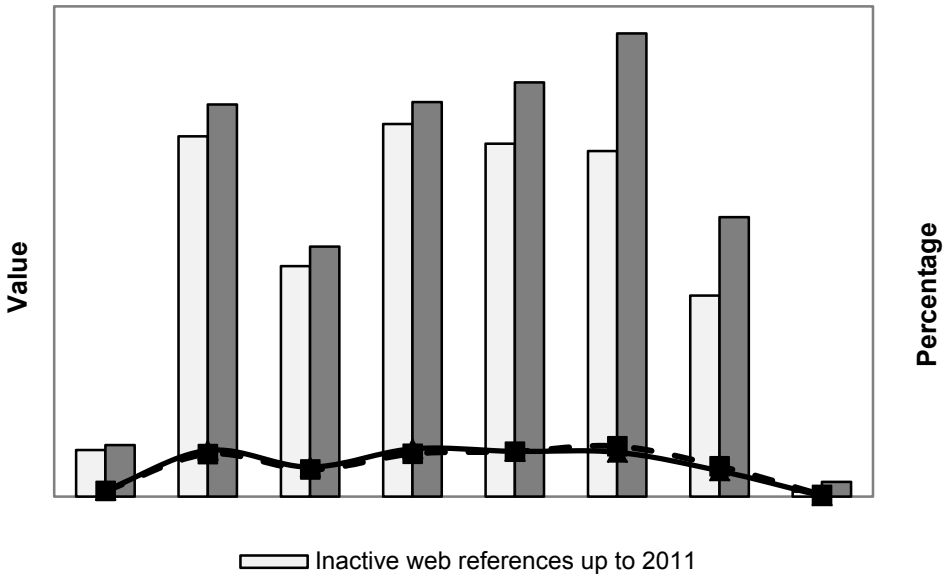


Figure 2. The change of inactive web references (from 2011 to 2013)

As depicted in Figure 2, there were more inactive than active web references in the period from 2003 to 2006, respectively, there were more active than inactive web references in the period from 2007 to 2010. Regarding the change in their availability, a higher number of previously active web references in the period from 2007 to 2010 have become inactive by the end of January 2013.
The premise was that the highest number of inactive web references is present in the first half of the analysed period. For instance, since the web references used in graduate papers in the year 2003 were accessed, until their availability was checked in the year 2013, ten years have passed, so one expects that most of them have disappeared.
The given assumption has been confirmed, i.e. there are more inactive than active web references from the first half of the analysed period, and also more web references that were active from the second half of the period became inactive in between the two studies, which indicates the rapid growth of inactive sources.

Inactive Web references according to the type of references

The data which was gathered for the second research (January 2013) was further analysed according to the type of inactive web references. The assessment of the type was based on the visual attributes of the web site where the reference was found (in some cases it was evident if there had been an article or some other electronic document on the web site). Another way to assess the type of references was possible by examining the bibliographic data provided in the graduate papers, and the type was identified with certainty for those references that had been active during the first research (June 2011).
If we categorise publication into monographs and periodicals, among 922 inactive web references there are 585 (63.4%) monographs, apropos 326 (35.4%) periodicals, while 11 (1.2%) of inactive web references were unidentifiable. Exactly 141 (15.3%) of inactive web references can be accurately identified by their type, since those are the ones that had been active by the end of the first research.

Table 1: Categorized types of monographs and periodicals (January 2013)

Monographs			**Periodicals**		
Type	**No.**	**%**	**Type**	**No.**	**%**
Web Pages	468	80.0	Articles on Web Pages	298	91.4
Documents /Databases on Web Pages	55	9.4	Scientific Articles	21	6.4
Legislation / Guidelines / Standards	41	7.0	Entries / Articles in Dictionaries	7	2.2
E-books / Manuals	15	2.6	-	-	
Lectures / Videos	6	1.0	-	-	
Total	**585**		**Total**	**326**	

The reason why there are so many monographs among inactive web references is that most of them are isolated web pages, which comprise 468, i.e. 80.0% of all monographs. On the other hand, articles on web pages comprise 298, i.e. 91.4% of all periodicals. The representation of the remaining inactive web references is shown in Table 1.
The assumption regarding the relation between the lifespan and the type of web references indicates that the isolated web pages have a shorter lifespan than any type of articles or other documents.
This was confirmed, as the majority of inactive web references are individual web pages, followed by the articles found on web pages. All other types are represented at a much lower rate.

Analysis of inactive Web references with respect to the type of Web sites
There was also a distinction regarding the types of web sites on which the inactive web references were found. The majority of them are too diverse in order to be classified into one group. They make 432 (46.8%) of all inactive web references, and most of them are commercial, non-scientific in nature. On the other hand, there are 263 (28.5%) web sites for the educational, cultural and scientific institutions while 223 (24.2%) web sites are considered to be undefined.
Interesting findings include the presence of four domains that were on sale, which is why the previous reference cannot be found on the given location. It was also noted that some of the frequently visited web sites have inactive web references, such as the web site for the National and University Library in Zagreb which included 32 inactive web references, the web site for *Narodne novine* (The Official Gazette) with 24 of them, and online encyclopaedia Wikipedia with 7 inactive web references. However, the majority of those references were relocated, but without a redirection from the former URL address which doesn't make them easily accessible.

It was expected that the majority of inactive web sites would be non-scientific, seldom visited by the users. The given hypothesis was partially confirmed since most of inactive web sites are commercial in their nature. However a vast number of web sites for educational, cultural and scientific institutions was not foreseen, as those web sites should be frequently updated and maintain the access to their references whether they are archived or not, since they are scientific in their nature.

The Presence of Authors and Publishers among inactive Web references
When it comes to the presence of authors and / or publishers among inactive web references, one expects that most of them do not have the aforementioned bibliographic data listed, which would be one of the reasons why they have a shorter lifespan.

As was anticipated, majority of the web references do not have an author, exactly 657 (71.3%) of them, and even more do not have a publisher listed, that is 848 (92.0%) of inactive web references.

Possible Solutions

There is a whole range of possible solutions to how to slow down the process of the disappearance of the scientific literature found online. Firstly, one can improve the citing system by implementing stricter, more formal obligation, and also, during the process of placing the references online, one should be required to list all the necessary bibliographical data.[7] As well as that, it is suggested to list more meta-data so that the search engines can index them more frequently and comprehensively.[8]

Thus far, the least complicated would be to archive outdated sources, that is simply relocate the reference, at the same time making it just as accessible through a redirection (the user is led to another location for the reference).[9] Also, there is a need to evaluate the usability of formats in which the references are accessible, so as to prevent them becoming obsolete,[10] just as one should take into consideration the design of the web site where the reference was found, such as its functionality and attendance rate.[11]

The results of this research correspond with the given findings, as well as outline one other possibility such as (whenever possible) citing more URL addresses for the same reference, or at least listing more references for the same information, so that in case one of them is no longer available, the data itself does not lose its validity.

Other possible solutions is adding extra identificator such as permalink for dynamic web sites or the digital object identifier (DOI) for the publications published by publishing houses. Both of this solutions alow more stable linking of online document. Because, in the case of DOI the publisher only need to change metadata for the DOI when URL is changed; and in the case of permalink even if context is replaced by something new, wanted contexts is still available with the same permalink.

[7] Maharana, B.; Nayak, K.; Sahu, N.K. (2006), p. 600.

[8] Spinellis, D. (2003), p. 74.

[9] Davis, R. M. (2010.), p. 2.

[10] Stanescu, A. (2005), p. 62.

[11] Shelstad, M. 3(2005), p. 209.

Conclusion

The main hypothesis, regarding the increase in inactive web references, has been confirmed, including the rate in which the web references have become unavailable in the given period. As well as that, the assumptions having to do with the type of web references, where web pages prevail, and the lack of authors and publishers among inactive web references, have also been confirmed. The only premise that has not been confirmed, in full, has to do with the type of web sites where the inactive web references were located, due to the fact that a high percentage of web sites were for the educational, cultural and scientific institutions.

Based on the results of the research and along with the previously stated findings, the reasons for shortening the lifespan of web references and the suggestions on how to increase it, have been summed up. The largest emphasis has been placed on the problem of unavailability of relevant scientific sources by invoking the need for stricter citing systems and laying out more formal obligations while placing the sources online in order to elongate their lifespan. The other main concern had to do with the inability to accurately estimate the average lifespan because of inability to get a hold of the exact date of the disappearance of a web reference. As a way of solving that, more frequent indexing is suggested. In correspondence to that, a more quality archiving system is proposed as well as evaluating the state of the formats for the references and the functionality of the environment in which they are found. One more solution had to do with listing more URL addresses or other identifiers for the same reference. Other possible solutions is adding extra identificator such as permalink for dynamic web sites or the digital object identifier (DOI) for the publications published by publishing houses, because both of this solutions alow more stable linking of online document.

To conclude, by getting insight into possible solutions on how to increase the lifespan of web references, and what are some of the reasons why they decay, one has a better understanding of what to do to insure their availability.

References

Ashenfelder, M. The Average Lifespan of a Webpage. Library of Congress. 2011. URL: http://blogs.loc.gov/digitalpreservation/2011/11/the-average-lifespan-of-a-webpage/ (11.07.2012).

Bauer, A.; Kirin, M.; Turković, M. Web References in Graduate Papers at the Department of Information Sciences at the Faculty of Humanities and Social Sciences in Zagreb. // *3rd International Conference "The Future of Information Sciences: INFuture2011 – Information Sciences and e-Society"*/ Billenness, C.; Hemera, A.; Mateljan, V.; Banek Zorica, M.; Stančić, H.; Seljan, S. (ed.). Zagreb : Department of Information Science, Faculty of Humanities and Social Sciences, University of, 2011, p.p. 69-75.

Coates, T. On Permalinks and Paradigms… 2003. URL: http://plasticbag.org/archives/2003/06/on_permalinks_and_paradigms (24.09.2013).

Davis, R. M. Moving Targets: Web Preservation and Reference Management. // *Ariadne*. The Web version (2010); 62. URL: http://pubs.ulcc.ac.uk/155/2/ariadne-print-issue62-davis.pdf (11.07.2012).

Franklin, J. Open access to scientific and technical information: the state of the art. // *Open Access to Scientific and Technical Information: State of the Art and Future Trends* / Gruttemeier, H.; Mahon, B. (ed.). Amsterdam : IOS Press, 2003. p. 74. URL: http://books.google.hr/books?id=2X3gW1lUvN4C&pg=PA74&hl=en#v=onepage&q&f=false (24.09.2013).

Halpin, H. Sense and Reference on the Web. // *Minds & Machines*. 2(2011), 21; p.p. 153-178. URL: http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=26&hid=110&sid=20746051-f82f-44a7-81e1-383db21aa97d%40sessionmgr112 (03.05.2012).

Maharana, B.; Nayak, K.; Sahu, N.K. Scholarly use of web resources in LIS research: a citation analysis. // *Library Review*. 9(2006), 55; p.p. 598 – 607. URL: http://dx.doi.org/10.1108/00242530610706789 (09.05.2012).

Meyer zu Eissen, S.; Stein, B. Genre Classification of Web Pages: User Study and Feasibility Analysis. 2004. URL: http://www.uni-weimar.de/medien/webis/publications/papers/stein_2004c.pdf (11.07.2012).

Online Extra: How the "Digital Object Identifier" Works. *BusinessWeek*. 2001. URL: http://www.businessweek.com/printer/articles/150904-online-extra-how-the-digital-object-identifier-works?type=old_article (24.09.2013).

Santini, M. Characterizing Genres of Web Pages. // *Proceedings of the 40th Hawaii International Conference on System Sciences* (2007) / R. H. Sprague, Jr. (ed.). Waikoloa: IEEE Computer Society. URL: http://www.google.hr/url?sa=t&rct=j&q=web%20page%20genres&source=web&cd=5&ved=0CF8QFjAE&url=http%3A%2F%2Fciteseerx.ist.psu.edu%2Fviewdoc%2Fdownload%3Fdoi%3D10.1.1.106.1160%26rep%3Drep1%26type%3Dpdf&ei=rKz9T5CnNfDN4QTGkvjdBg&usg=AFQjCNFz0ANLN7DD_9IpTwL5xN7Wj3Ndrg&cad=rja (11.07.2012.).

Shelstad, M. Content matters: analysis of a website redesign. // *OCLC Systems & Services*. 3(2005), 21; p.p. 209-225. URL: http://dx.doi.org/10.1108/10650750510612407 (09.05.2012).

Spinellis, D. The Decay and Failures of Web References. // *Communications of the ACM*. 1(2003), 46; p.p. 71-77. URL: http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=11&hid=110&sid=3d002ef8-7047-449f-83e9-262cf1cc8873%40sessionmgr104 (03.05.2012.).

Stanescu, A. Assessing the durability of formats in a digital preservation environment: The INFORM methodology. // *OCLC Systems & Services*. 1(2005), 21; p.p. 61 – 81. URL: http://dx.doi.org/10.1108/10650750510578163 (09.05.2012).

What is the Average Lifespan of a Web Site. DialMe.com, Webmaster Tips and Resources. URL: http://www.dialme.com/blog/what-is-the-average-lifespan-of-a-web-site/ (11.07.2012).

A Domain-Specific Records Management and Information Governance Solution Designed to Support the Implementation of the General International Standard Archival Description

Arian Rajh
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
rarian_sp@yahoo.com

Krešimir Meze
Omega Software Ltd.
Oreškovićeva ulica 25, Zagreb, Croatia
kresimir.meze@omega-software.hr

Summary

The authors investigate the creation of a records management (RM)-and-archival system based on the General International Standard Archival Description standard (ISAD(G)) and domain specific language (DSL) with RM and archival functions already built-in. This should enable harmonised archival descriptions and better communication between RM subject-experts and developer experts. This should also shift the focus onto solving the creators' needs related to their instances of RM processes and the characteristics of their archival material. The developed RM-and-archival system solution aims to streamline RM business processes and support them through the implementation of specific automation procedures, information technologies like barcode/RFID and additional controls. This RM-and-archival system supports the process of creating a records schedule, a procedure for registering archival material, the retrieval and usage process, storage management, ISAD(G) compliant archival description, disposal processes, inventory checks and the production of XML records with information about archival material. The authors conclude by announcing future work on aligning RM and archival DSL with the MoReq2010 set of services.

Key words: Archival Description, Domain specific language, ISAD (G), Records Management and Information Governance

Introduction

Public and private organisations have to manage large amounts of records that they receive and create while running their business processes. Record management (RM) practice in public organisations usually differs from RM practice in private organisations. Public institutions receive a large number of applications from citizens, private organisations and other public organisations destined for long-term preservation. Public institutions usually have highly demanding workflows and case management procedures, large quantities of additional support documentation and a large number of different types of documents. Processes that circulate documents between working colleagues are not just support processes in the public sector; they have high significance for successful case management in public institutions.[1] In addition, public institutions are under the auspices of archival authorities and they have to manage their records in a way that meets the regulatory environment, mandatory and voluntary practices and archival instructions.[2] The design of records management practice and the selection of RM or DMS/ERMS (Document Management System/Electronic Records Management System) tools have become very important for public institutions.

Successful RM practice should cover all records-related tasks from receiving them to archiving them according to previously developed rules. RM and archiving in the creators’ organisations is a domain that comprises registration, storage and arrangement, description, disposal and cooperation with the archival authority. These processes have their own logic and rules. The functions, concepts, terms and business roles in these processes are carrying a specific RM and archival meaning. In order to build a tool that supports RM and archival business processes, it was presumed that it was necessary to build use cases of the processes and a model of the tool using a domain specific language (DSL). RM and archival DSL were built using terms and concepts gathered from professional archival experience and further developed on specialised platform for DSL and system development (the Rhetos platform). This platform generated code, database schemes, a user interface and web services for RM-and-archival system in development. Using Rhetos boosted the development of the system, decreased the misuse of functions and created possibilities for the permanent updating of the newly-constructed DSL. The final result of this development was the RM-and-archive system for the public sector (at first as a prototype and later as a final product Centrix).

[1] Brumec, Dobrović, Tomičić: The Model of the DMS in the Public Sector. Journal of Information and Organizational Sciences, Vol.30, No.1, 2006, page 32

[2] ISO 15489-1:2001, p. 4-5.

Organisational and process-related transformation

Documents and archival records are used and reused in the business processes of public institutions as a basic input of these processes, their result or their supporting material. Since records managers often lack certain descriptors for case- and business-oriented retrievals at the moment of registering or describing archival units, RM and archival system was built as a modular system that can communicate with main business registries via an intermediary. This is consistent with the MoReq2011 concept that endorses the usage of main business applications and RM functionalities.

The RM-and-archival system was developed upon a precise professional DSL and it was made as completely process-oriented software. The main processes used as the backbone during the development were receiving and registering technical units, filing or placement of technical units in storage, archival description, distribution and usage of material, preservation/disposal and collaboration with the archival authority. The RM and archival processes were streamlined during the development of the tool, and their tasks were reengineered in order to support them with information technology. These streamlined processes can be used as an instrument of organisational transformation in institutions with insufficient RM practice.

Defining the retention schedule, adding an archival description and the process of registering the technical units

RM and archival material consist of various technical units, i.e. conventional units like binders or unconventional units like optical disks etc. Technical units are organised into RM/archival units according to the records schedule. The records schedule lists groups and items of records, their descriptions and information on the retention period. It is linked with information on the level of archival description for a particular RM/archival unit, types of retention period start date[3] and permanent preservation/disposal procedures. The retention schedule is a hierarchical classification scheme that governs RM/archival units with the technical units that belongs to them. If a retention schedule is linked to a case management tool, through a case classification number, with accurate metadata on the retention period start date for particular case, and if the technical units are linked to this case, the records schedule should trigger the automation of the retention procedure. The retention schedule could also be linked with the DMS/EDRM and govern archival storage procedures.

[3] After finishing a case, at the end of the year in which the unit has been created etc.

Table 1: A records schedule extended with metadata for the automation of the retention procedure

Ordinal number	Title	Description	Description level	RP	RP start date	Disposal procedure
n.n.	group	description of the group	series	-	-	-
n.n.n.	item	description of the item	subseries	x years	A	destruction
n.n.n.	item	description of the item	subseries	x years	B	preservation

There is some additional metadata related to this kind of grouping of records. Conditions of access could be added to each item of the schedule. When this is combined with the user management service and ascribed to the workflows of the DMS or business registry/applications, it could help the system to govern the access function and related interfaces.[4] Creators could also use this master data to create an information catalogue etc. Records schedules should be customised by the creators' records manager. After defining the records schedule, it is possible to perform the process of archival description and to define the list of RM/archival units. The list of RM/archival unit contains information on existing items in the records schedule. Items in the list represent real RM/archival units and items in the records schedule represent the set of potential units for one creator (and rules for their preservation). Besides providing detailed information for RM/archival units, archival description fills the gap between usually obsolete content-based retention schedules and the contemporary RM/archival units' hierarchies needed for the function-(or process)-based management of records. The functional organisation of records represents the principle of provenance and links together archival holdings, functions and creators. Archival description metadata is gathered by the areas of description compatible with ISAD(G). By adding an ISAD(G)-based archival description, an RM-and-archival system gains an archival search tool. The process of archival description incorporated into the solution follows the multilevel description rule: metadata depends on the level of description, all descriptions are linked together, and the metadata is not repeated through different levels. Although it is intended that information on more general description units (fond, series) will be entered first and then specific units are described, records managers can always add additional information to hierarchically super-ordinated units. Areas of description are: identity statement area, context area, content and structure area, conditions of access and use, allied materials area, notes area and description control area.[5] The integra-

[4] Reference model for an open archival information system (OAIS), Recommended practice CCSDS 650.0-M-2, Magenta book, June 2012, p. 4-16 (part 4.1.1.7), http://public.ccsds.org/publications/archive/650x0m2.pdf, accessed in March 2013

[5] The identity statement area in a Centrix RM-and-archival system comprises identifiers, superordinated units, level of description, unit title, additional title and dates. The context area com-

tion of the ISAD(G) attuned search tool into this system improves the creators' archival management functionality, it bridges the domains of the records manager and archivist and provides a tool for the real archival organisation of emerging material. This system links a records entrance point and archival point (links between technical units and cases from the case management tool). It also facilitates collaboration between creator and archival authority using XML schemes for export of data on archival holdings to the national archival system.

The process of the registration of technical units in the prototyped and developed tool is designed to be flexible enough to enable an optimal workload for records managers. They can choose to register all the technical units if necessary (e.g. if tasks are executed simultaneously and technical units linked to the same case are used by different employees). In the cases of simpler workflows, they can choose to register larger technical units – it depends on the creator's actual RM practice. Registration of technical units starts when an office employee submits technical unit(s) into the archive. The records manager receives it, links it with the item of records schedule and enters metadata. Metadata categories are related to XML schemes of the ARHiNET – national archival information system. The record manager enters the title of the technical unit, year of creation, type and quantity of technical unit, medium/media, preservation-related metadata and a note. Records manager selects an item of the records schedule the technical unit is related to. The system fills in the RP, RP start date, RP end date, submission to archival authority date, availability rules, identifiers and date of entry.[6] The records manager links the technical unit with the item of the existing RM/archival units list. System then gathers and adds quantitative metadata from technical units to registration/archival unit's level and this can be used in an XML export for ARHiNET. The final task of the registration process is to create a label for the technical unit. The label contains the required data but it can also carry a barcode and RFID tag if warehouse management is supported.

---

prises the name of the creator and its administrative/biographical history, the archival history of the unit, and the source of acquisition/transfer. In the content and structure area is metadata for: scope and content; appraisal, destruction and scheduling information; foreseen additions to the unit; and system of arrangement. The conditions and access of use area lists information on conditions of access, conditions of reproduction, language and scripts, physical characteristic and technical requirements, existing search aids/registration office aids. In the allied materials area, the record manager should provide users with information on the existence of originals (if the unit consists of copies), the existence of copies (if the unit has been reproduced), related units of description, and bibliographies/publications. The notes and description control area consist of a note, a record manager/archivist note, rules and conventions used for creating the description of the unit, and the date of the description/revision of the description. In addition, archival description functionality also provides users with a list of technical units that belong to a particular RM/archival unit.

[6] This basic set of metadata could be extended, which depends on archival material types and the creator's RM requirements. Describing technical documentation, cartographic material, photographs, microfilm, AV material or electronic records requires specific metadata.

Filing process

The filing process should have the basic characteristic and functionalities of a warehouse management system so it can be supported by imager or radio frequency technology device. Records managers should define the hierarchical structure of the creator's storage. When technical units are placed in some organisational unit's storage or in the archive (with the status: archived), records manager pairs location information with the technical unit. Filling process could be enhanced by using barcode or RFID terminals for reading the barcode of the position and pairing it with the barcodes or tags of technical units, and by using sensors and RFID readers placed beside the entrance of the archive port, for controlling what enters or leaves the archive.[7]

The distribution and usage of archival material

When a user retrieves material and sends an order, the record manager finds technical units in the archival storage and takes them back to the shelf after usage. Records management/archive system should track the orders for archival material and their deadlines. System can be customised to synchronise users' orders with taking archival material from the storage (with an RFID port reader and sensors that activate it) and with filing forms for confirmation of taking over the materials by users (with barcode readers). The distribution and usage function is supported by extended retrieval possibilities. This should be done because the typical user is likely to use the case-related descriptor before the title of the technical unit that he requires. However, a typical record manager can't have entered case-related descriptors because they were usually not known. This is why retrieval possibilities are extended by enabling communication between the RM system, case management system and business registry by the use of an *intermediary* agent.

Permanent preservation/disposal process

The records schedule and archival description areas for the particular RM/archival unit contain temporal and event-related metadata required for calculating the retention period end dates (according to purpose-designed algorithms). Some technical units are intended for long term preservation in the creator's organisation and archival authority, and some are destined to be destroyed after the retention period ends. The disposal process starts with listing the technical units that should be destroyed. The system tracks down temporal metadata (date of decision, metadata on the start of the retention period, retention period), calculates the retention period end date and notifies the record manager. After that, the record manager asks the archival authority for approval

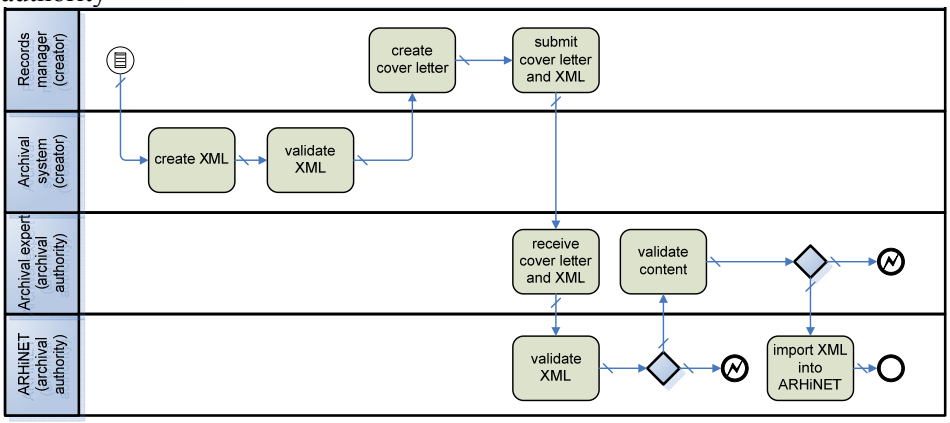
[7] A good example of applying RFID/barcode technologies is an RM project realised in the Croatian Agency for Medicinal Products and Medical Devices in 2013 (the project partners were Omega Software and Selmet).

for the destruction and updates information on what has been disposed of. For archival units that need to be preserved and transferred to the archival authority, the system prepares notifications for creator's records manager and a list for the archival authority (with a temporal trigger for the period set up by the archival legislation).

Collaboration with the archival authority

Legislation requires creators to inform their archival authorities about their RM/archival holdings once a year. This could be done in Croatia directly by using ARHiNET (national archival system) or by submitting an XML document with metadata on RM/archival holdings. This document should provide the archival authority with information on RM/archival units under the creator's care. The business process starts after the event trigger and includes the automatic creation and filing of an XML file,[8] sending a cover letter and the XML file by the creator, content-related and technical validation of the XML file by the archival authority, and the upload of metadata to the ARHiNET.

Business process model 1: The submission of data from the creator to archival authority



A streamlined business process of submitting data for creator's archival material is shown in the model below. Validation in the creator's organisation could be performed by RM/archival systems or XML editors. There are two possible improvements to the model. The first one is related to preventing the creation of invalid XML. The archival system in the creator's organisation could enable the validation of XML at the moment of creation by executing encoding rules. The second improvement is related to building a gateway for transferring XML from the creators' systems to ARHiNET directly.

[8] An XML file is created according to instruction on the ARHiNET web site: http://arhinet.arhiv.hr/_Pages/PreuzmiXMLShemu.aspx, downloaded in April 2013.

The development of the RM-and-archive system for the public sector

Record management is essential to all organizations. The ISO 15489 standard emphasizes that records represent evidence of a business transaction or decision and that they can be used in the pursuance of legal obligations. Despite significant differences in business processes, this applies to both public and private organisations. To fulfil evidential requirements, it is important to assure that every record in an organization is kept as authentic, reliable, integral and usable. It is obvious that an RM database for evidence and ERMS for archival storage can be very helpful in record management practice. In addition to these expectations, a state of the art system must satisfy even more challenges. Legal regulations and professional standards in specific business domains must be met. They usually require a high level of business knowledge from the system architect and developers. Also, each organization or company has its own specific record management needs so the system must be flexible and customizable in order to adapt to them. Due to rapid technological changes, ERMS must now be able to adjust to new trends in a reasonable time frame and within a reasonable cost. Standalone ERMS systems have existed on the market for some time, but customers in the first place focus on solutions that meet their core business requirements and prefer record management functionalities built in line with business applications. The main question is how to produce such a system efficiently?

The key concepts of the development based on the use of RM and archival DSL

The methods, tools and programing languages that are currently widely used in software development are still not immune to errors in the design phase and in performance. This makes IT solutions based on them difficult to develop and expensive to maintain. This is why engineers at *Omega Software* decided to take an approach based on the innovative use of domain-specific languages (DSL) for prototyping and building a RM and archive system. Domain specific language is a kind of language based on terms, keywords and rules originating from some problem-specific domain. DSLs standardize and simplify information exchange between actors in specific domain. Probably the best known example of a DSL term is the H_2O formula. This is a term written in chemical DSL. DSLs are not reserved for purely scientific domains though. They are widely used in applied professional and business areas. Experts in finances use keywords like *invoice* or *payment*. In the governmental domain, we can encounter *case* or *file* and in record management we use *record*, *aggregation* or *class*. These are all well-defined professional concepts with a clear and strict meaning and associated business rules, so they can also be used as concepts in programming languages. Languages based on concepts taken from specific domains are called domain specific programming languages. This means that programmers have abstract concepts for describing the models they want to build.

These concepts are already integrated into business domains so they could be used for information exchange between the programmer and future software users. Hence domain specific languages mainly consist of understandable functions and data types they could significantly simplify communication between development engineers and users. A cleverly designed programming language delivers one additional benefit: the main focus of the programmer is now shifted from *how to do something* to *what to do precisely*. Beside definition, the DSL concept brings certain functionality and business behaviour from the domain it pertains to, so strong expertise in business domain is no longer required for a software architect and developer. Concepts themselves carry behaviour compliant with business rules.

There are two prerequisites for the development of systems based on the above stated paradigm. Firstly, a domain specific programming language and its concepts should be designed. Programmers will later use these concepts to describe the actual business data model with entities, metadata and business rules in DSL scripts. DSL scripts are plain text files that are easy to read and edit in any text editor. Secondly, the DSL execution platform should be implemented. The execution platform is a software component that reads DSL scripts written by the programmer and generates server application, database and web services used for communication with other systems and applications. In particular, the application can be a web or mobile user interface for interaction with human users or another line of the business system. Communication between the application and server side web services happens through the REST (Representation State Transfer) or SOAP (Simple Object Access Protocol) interface. REST and SOAP are industrial standards for message exchange between IT systems and applications.[9] By separating the business logic to a server application, the consistency of data and the execution of business rules are obtained. Any number of client applications can then use services that the server application provides, and none of them need to implement any additional business logic. All data validation procedures, security and authorization rules are also executed in the server application. *Omega Software* has implemented its own execution platform called *Rhetos*. The *Rhetos* platform was constructed using several basic DSL principles. The main purpose of starting the project was the overall desire to improve the quality of the application design process, the solutions themselves and the achievement of long-term support (extension, adjustment to specific businesses, convergence with the general and specific norms of the domain etc.). Applications built on the *Rhetos* platform are not constrained to one particular domain or one particular domain specific language. *Rhetos* can be used for the creation and execution of domain specific languages for any domain and subdomain. This platform is extremely extensible, so programmers can add new concepts

[9] For additional information about these industrial standards, please see: http://www.w3.org/TR/soap/ and http://www.w3.org/TR/ws-arch/ (accessed June 2013).

tied to any business domain. The whole internal infrastructure (e.g. database management system) can be changed in *Rhetos* from the default to any other required by the customer.
Rhetos platform is published as an open source platform. This enables users to fully grasp the content of the platform and to participate in design and changes. The open source *Rhetos* platform brings several meaningful benefits to implementers and users. Most significant is reducing the price of development and facilitating community support. Furthermore, it increases security and user awareness of the organization and implementation of corrections and upgrades, without the need to consult the original implementer of the software solution.
The next step in development will be standardization with MoReq2010, the latest industrial standard for DMS/ERMS. This was written by a pan-European community interested in Information Governance across Europe. MoReq2010 aims to provide a comprehensive but easily understood set of requirements for an RM system that is intended to be adaptable and applicable to divergent information and business activities, industry sectors and types of organization. MoReq2010 is the third version of the MoReq specification and it enables the certification of business applications with standardized RM functionalities, in addition to the certification of standalone ERMS systems. In the future, it is planned to implement the MoReq2010 services and concepts as a particular DSL that will be used for extending other additional business DSLs with harmonized RM functionalities. This will also bring benefits to independent software vendors, especially SMS companies. Although the MoReq2010 standard is written to be simple, implementing the MoReq2010 requirements is a complex and expensive task to do. The MoReq-based DSL will be a useful extension for providing business solutions with standardized record management functionalities.

Conclusion

The new approach used for the development of an RM-and-archive system included the creation of model and system by using RM/archival DSL for most domain-specific processes and by using the ISAD(G) international standard for the implementation of archival descriptions. During this development process, it was shown that using DSLs enables the creation of systems that already contain domain-related professional logic. The usage of ISAD(G) enables the creation of description of archival material according to immanent archival instructions. Finally, Moreq2010 compliant records management domain specific languages could be the basis for the creation of records management and archival systems that would fit a large variety of organisations.

References

ARHiNET Croatia, http://arhinet.arhiv.hr (April 2013)

Brumec, Dobrović, Tomičić: The Model of the DMS in the Public Sector. Journal of Information and Organizational Sciences, Vol.30, No.1, 2006

ISAD(G): General International Standard Archival Description, 2nd ed.

ISO 15489-1:2001 Information and documentation - Records management - Part 1: General

ISO 14589-2:2001 Information and documentation - Records management - Part 2: Guidelines

MoReq2010 Modular Requirements for Records Systems, http://moreq2010.eu/ (April 2013)

Reference model for an open archival information system (OAIS), Recommended practice CCSDS 650.0-M-2, Magenta book, June 2012, http://public.ccsds.org/publications/archive/650x0m2.pdf (March 2013)

World Wide Web Consortium (W3C), http://www.w3.org (June 2013)

The InterPARES Trust Project – Trust and Digital Records in an Increasingly Networked Society

Adam Jansen
School of Library, Archival and Information Studies
University of British Columbia
470-1961 East Mall, Vancouver, BC, V6T 1Z1, Canada
adamj@mail.ubc.ca

Luciana Duranti
School of Library, Archival and Information Studies
University of British Columbia
470-1961 East Mall, Vancouver, BC, V6T 1Z1, Canada
Luciana.Duranti@mail.ubc.ca

Summary

Key issues of ownership, jurisdiction, and privacy regarding Internet-based records have yet to be resolved, but technology will not stand still to wait for legal and regulatory systems to catch up. There exists a need for a shared regulatory and procedural framework that promotes consistency and balance in terms of policies and practices regarding the handling of digital objects. The InterPARES Trust Project will generate new knowledge on digital records maintained online and will create a new model statute specific to the Internet that will enable legislative reform work for the continued development of the current fleet of uniform statutes.

Key words: Trust; Digital Records; Cloud Computing; Social Media; Privacy; Access; Secrecy; Digital Preservation; InterPARES

Introduction

Advances in the speed and connectivity of computer resources have allowed individuals to create, store and access vast amounts of records in the highly networked, and some would argue easily hacked, environment of the Internet. Organizations such as banks, public utilities, hospitals and governments which are trusted to store the often personal and sensitive records of these individuals are also increasingly utilizing Internet storage and/or providing Internet access to records without fully understanding the implications. Where these records are being stored, how they are being managed, and how long these records will remain accessible is unclear. As Internet resources can, and often do, reside across juridical boundaries, who is liable for any security breaches and maleficence is unclear and most Internet storage providers' service level agreements (SLAs)

exempt these providers from prosecution.[1] The accumulation of massive amounts of data in order to provide a host of services, many of which are focused on marketing and obtaining competitive advantage, is increasingly centralizing personal information in 'big data' warehouses. These warehouses are ripe with valuable information and attractive targets to computer criminal activity.

The issues that arise from such an increasingly networked world of data are clear: Can the records be trusted? Can records that derive from data stored online be considered trustworthy or even traceable? Are they complete? Are they authentic? Who created them, modified them, and access them? Where are they stored and under whose jurisdiction would legal disputes be heard? Who is securing the systems, and how can it be proven that the records remains safe and unaltered?

Goals and Objectives

The goal of the InterPARES Trust Project is to generate the theoretical and methodological framework that will support the development of integrated and consistent networks of policies, procedures, regulations, standards and legislation that can be applied to digital records that are created, stored and accessed using the Internet. These frameworks can increase and validate the public trust by providing evidence of good governance, a strong digital economy and a persistent memory. To achieve this goal, the objectives of the research are to:

1. discover how current policies and practices regarding the handling of digital records by institutions and professionals affect the public's trust in them, in light of the exponential growth of and reliance on Internet services;
2. anticipate problems in maintaining any trust in digital records under the control of entities suffering a waning level of confidence from the public (including legal, law enforcement, financial, medical, broadcasting, "hacktivist," and governmental organizations and professionals);
3. establish what significance national/cultural contexts have with regard to the level of trust digital records on the Internet enjoy;
4. articulate model policies, procedures, and practices for creating, managing, accessing, and/or storing records on the Internet, especially in social media and cloud computing environments and through mobile technolo-

[1] As an example, the Amazon Web Services Customer Agreement states: "The service offerings are provided 'As Is.' We and our affiliates and Licensors make no representations or warranties of any kind, whether express, implied, statutory or otherwise regarding the service offerings or the third party content, including any warranty that the service offerings or third party content will be uninterrupted, error free or free of harmful components, or that any content, including your content or the third party content, will be secure or not otherwise lost or damaged.." This agreement can be found at http://aws.amazon.com/agreement/

gies, and test them in a variety of contexts so that, from them, international standards, guidelines and best practices can be developed; and,

5. formulate proposals and models for law reform, and functional requirements for the systems in which Internet providers store and manage digital records.

Theoretical Framework

The theoretical framework used for this project is being adapted from archival and diplomatic theory, with particular emphasis on the concepts that are foundational to trusting records.[2] The research will use resource-based theory, which focuses on the importance of the technical, managerial and relational capabilities of an organization for leveraging available resources to maximize competitive advantage.[3] Resource-based theory illustrates the performance differences between organizations by analysing the way in which they leverage the resources available. Through utilizing resource-based theory, resources that are unique to cultures, societies and various types of organizations can be identified and articulated into a global model.

Risk management concepts will also be incorporated in the framework by analyzing the question of trust in relation to the amount of security, risk and exposure that an organization is willing to accept to achieve its goals. Current research and practice in the field of risk management offer both an operational and a social perspective on trust. Operationally, risk management concepts[4] have been increasingly applied to archival science to expand the understanding of risk beyond the loss of the primary object, or record, to also encompass threats, vulnerabilities and mitigations. Of particular importance to the survivability of records is the threat of technological obsolescence that was investigated in prior phases of the InterPARES project.[5] Socially, factors that contribute to the establishment or eventual erosion of trust are of interest to the study of risk management. The public is asked to trust governments and organizations while simultaneously being exposed to increasing evidence posted on the Internet detailing corruption, scandal and large-scale environment disasters.

Lastly, to design effective model policies that are applicable across a broad spectrum of organizations, cultures and societies, the project will draw upon design theory. As the policies developed will need to address challenges that arise from interaction with future technologies not yet imagined, the project will need to adopt an "argumentative process in the course of which an image of the

[2] MacNeil, 2000; Duranti and Preston, 2008

[3] Alvarez and Barney,2002

[4] Such as *IS0 31000 Risk Management – Principles and Guidelines on Implementation.*

[5] Duranti and Preston, 2008

problem and of the solution emerges gradually among the participants, as a product of incessant judgment, subjected to critical argument."[6] Thus the project will begin by taking into account design perspectives, not just in terms of how to utilize the theory, but through direct interactions with the developers of digital information technologies. A particular focus of will be on Human-Computer Interaction (HCI), as researchers and practitioners of this field design the interfaces that are used to create, store and access records over the Internet.

Methodological Approach

Research of this nature necessitates a multi-faceted, diversified, and dynamic approach: multi-faceted in order to deal with the polymorphism of digital records produced and accessed using a myriad of digital devices; diversified in order to accommodate the varied requirements produced by different social, cultural, and organizational contexts; and dynamic in order to respond to the rapid pace of technological change, while accounting for current uses as well as future uses and expectations.[7] The project will employ a research methodology based on empirical observation (through case studies) generated through an ethnographic approach and surveys. To understand concepts like trust in the case studies, the project will use the general theory of social construction, building hypotheses upon a foundation of the findings from previous InterPARES research.[8] The project will focus on gathering, analysing, and interpreting data gathered from a diverse cross-section of organizations and institutions from across the globe in order to explore the nature of the trust relationships that exist between Internet connected parties, as well as the risks, weakness and faultlines that are inherent in the storage and management of the records that exist 'in the cloud.'

The case studies and surveys conducted for this project will be carried out over a four year period beginning June 2013. At the conclusion of each study, the findings will be represented using activity and event models designed to provide an understanding of the situational realities and work processes that exist at each site both before and after changes suggested by the research team. Each modelling activity will be accompanied by diplomatic and archival analysis, digital forensic analysis, textual analysis, and visual analytics. Finally, comparative analysis will be used to generate a theory of what constitutes trust in networked environments that is capable of transcending juridical boundaries.

[6] Ritter & Webber, 1973, p. 162

[7] Thibodeau, 2012

[8] http://www.interpares.org

Project Deliverables

This project will generate new knowledge regarding digital records maintained on Internet connected resources and accessed through a myriad of digital devices. Such knowledge, developed by an inter-disciplinary team of international scholars and professionals, will produce new methods for identifying and protecting the balance between privacy and access, secrecy and transparency, the right to know and the right to forget. These methods can be used as the foundation for model legislation related to e-evidence, cybercrime, security and privacy. The research outcomes will be central to developing methods for the authentication of identity on the Internet and for protection against on-line fraud. By enhancing identity management, network security is also increased. Finally, the results will provide a sound basis for developing policy models, procedures and standards to manage records on the Internet; education modules for professionals and academic curricula for graduate programs; functional requirements and specifications for securing online records systems; and analytic frameworks that can be used to evaluate business models emerging from and only possible in the evolving Internet environment.

Collaborative, International Research Team

The team of researchers working on this project comprise of universities and organizations, both public and private, from around the world with expertise in the disciplines of archival science, records management, diplomatics, law, policy studies, health informatics, journalism, information governance and assurance, computer science, cyber-security, and digital forensics. As an international, collaborative project, the teams are organized into five major groups: North American, Latin America, Europe, Asia and multinational organizations. Among the larger participating institutions are: National Archives of Mexico, National Archives of Brazil, British Library, European Commission Anti-Fraud Office, International Federation of Red Cross and Red Crescent Societies, International Monetary Fund, International Records Management Trust, Israel State Archives, National Institute of Standards and Technology, NATO, Renmin University of China, Mid-Sweden University, University of British Columbia, Government of British Columbia, State Archives Belgium, UNESCO, University College London, and University of Washington. Croatia is strongly represented in the European group with a research team from the University of Zagreb – Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences (FHSS) and partners from the Croatian State Archives, the Faculty of Organisation and Informatics, the National and University Library, the Croatian Information and Documentation Referral Agency of the Government of the Republic of Croatia, FINA, the University of Zagreb – University Computing Centre, and Teched Consulting Services, Ltd.

Conclusion

The goal of the project is to generate the theoretical and methodological frameworks that will support the development of an integrated network of policies, procedures, regulations standards and legislation that can be consistently applied across juridical boundaries in order to ensure public trust grounded on evidence of good governance, a strong digital economy, and a persistent digital memory. This project will focus on the relationship between organizations (both public and private) and their client groups (i.e. citizens, customers, students, etc.) and the degree of trust that clients can place on the records of these organizations that are created, stored or access from the Internet, as well as the level to which organizations are concerned about establishing and maintaining a trust relationship with their clients. This focus on the trust relationship will develop new knowledge regarding records kept on social media and in the cloud and provide new methods for determining the appropriate balance between privacy and access, secrecy and transparency, the right to know and the right to be forgotten in globally connected networks. The research will propose changes to existing legislation and infrastructure, develop model policies, procedures, and practices to store and manage digital records over the Internet, and create functional requirements for networked systems utilized by Internet providers.

References

Alvarez, S.A., Barney, J.B. (2002). "Resource-based theory and the Entrepreneurial Firm." In Hitt, M., Ireland, R., Camp, S., Sexton, D., eds., *Creating a New Mindset: Integrating Strategy and Entrepreneurship Perspectives.* John Wiley and Sons Inc., New York.

Amazon Web Services, Inc. *AWS Customer Agreement*. 15 March 2012. Web. http://aws.amazon.com/agreement/. 28 May 2013.

Duranti, L. and Preston R., eds. (2008). *InterPARES 2: Interactive, Dynamic and Experiential Records.* Padova: ANAI.

International Research on Permanent Authentic Records in Electronic Systems (InterPARES). *The InterPARES Project.* Web. http://www.interpares.org. 28 May 2013.

International Standards Organization. (2009). *ISO 31000:2009. Risk Management – Principles and Guidelines on Implementation.*

MacNeil, H. (2000). *Trusting records: legal, historical, and diplomatic perspectives.* Kluwer Academic Publishers, The Netherlands.

Rittel, H., and Webber, M. (1973). "Dilemmas in a General Theory of Planning," *Policy Sciences*, Vol. 4, Elsevier Scientific Publishing Company, Inc., Amsterdam: 155–169. [Reprinted in Cross, N. (ed.). (1984) *Developments in Design Methodology*, J. Wiley & Sons, Chichester: 135–144.]

Thibodeau, K. (2012). "Wrestling with Shape-Shifters: Perspectives on Preserving Memory in the Digital Age." In *Proceedings of 'The Memory of the World in the Digital Age: Digitization and Digital Preservation'*, forthcoming.

Sense-making for Records Managers: Taking a Leading Role in SharePoint Governance

Nermina Bogičević*
Special Tribunal for Lebanon
PO Box 115, 2260 AC Leidschendam, The Netherlands
bogicevicnm@gmail.com

Summary

This case study of four SharePoint implementations within a single organization shows that applying sense-making framework Cynefin can help in good and bad times during project and may even anticipate surprising user-influenced ultimate solutions. Role of records managers in these implementations is elaborated and suggestions are made on how to improve their role in such projects and in organizations in general.

Key words: SharePoint Governance, Electronic Records Management, ERM, Cynefin, Project Management

Introduction

In the constant search of the "best practice" a records manager may find him or herself swimming in the infinite sea of information on the new paradigms and the frameworks and tools to tame these paradigms. One may find that the "records continuum" theory might be exactly the consoling answer to organizational chaotic electronic records management (ERM) while taking the agile approach to software development might just solve the impossibly complex arena of stakeholders wants and realistic budgetary possibilities. One particular article that has named ERM a "wicked" problem, thus provided comfort to those struggling and brought people issues into this arena up front (Childs and McLeod, 2013) has served as inspiration for this case study of four SharePoint implementations within a same organization. Whether SharePoint may appropriately address proper ERM is not the subject of this article but the different approaches to final solutions that may also evolve as an unplanned and unexpected result of same or similar implementations. In addition the role of records managers in the implementations will be highlighted.

* The views expressed herein are those of the author and do not necessarily reflect the views of the STL.

Projects Background

The four SharePoint implementations were all coming from different organizational offices and were all requested by the users themselves. These were essentially all related to improvement of records management and thus required organizational records managers involvement and guidance. Detailed description of projects' nature is described in Table 1. along with information on whether a project team was assigned or not and finally the projects' duration. As records in question did not belong to the primary category of this organization's records directly related to its primary function but belonged to other two categories (records generated in relation to the primary function or administrative records category) SharePoint as organizational secondary ERM system was recommended while the long duration of projects can also be explained by the fact that these projects were not related to organization's primary functional priorities. Only in Project A was there determination to successfully close the project within reasonable time as records originated from one of the most important organizational offices and records registration depended on the project's success.

Table 1.

Project	Nature of Project	Project Team	Project Duration
A	Improvement of records management and collaboration	Yes	9 months
B	Creation of intranet presence; Improvement of records management and collaboration	Yes/No	6+ months (on-going)
C	Improvement of records management and business processes	No	6+ months (on-going)
D	Improvement of records management	Yes	6+ months (on-going)

Role of Records Managers in Projects

It is worth mentioning that Project A was initiated after records management short-comings were identified during organization-wide records survey that resulted in Content and Data Management Strategy[1] recommending (among others) targeted use of SharePoint for management of records with limited retention periods. Although records in Project A had longer or permanent retention periods the use of organizational primary ERM system for their management was not possible in due time, thus the SharePoint solution had to be tailored with export to another system in mind. In-place records management was chosen and due consideration was given to records file naming. In the process organization-wide content type and taxonomy controls were developed of which many other SharePoint projects benefited. This undertaking has also added time

[1] Special Tribunal for Lebanon. Content & Data Management Strategy: Ensuring the STL Legacy, June 2012

needed to complete the project. All other projects were requested specifically by the user offices and were coordinated by a SharePoint working group. Project Team for Project A consisted of users, information technology representatives and was led by records managers. During the implementation 65 different issues were identified. Table 2. shows the breakdown of whom the resolution of these issues was allocated:

Table 2. Number of allocated project issues by stakeholder in Project A; Records managers (RM) Information Technology staff (IT) and users.

Stakeholder	Number of Issues
RM	24
IT	6
RM and IT	4
RM and user	18
IT and user	1
User	8

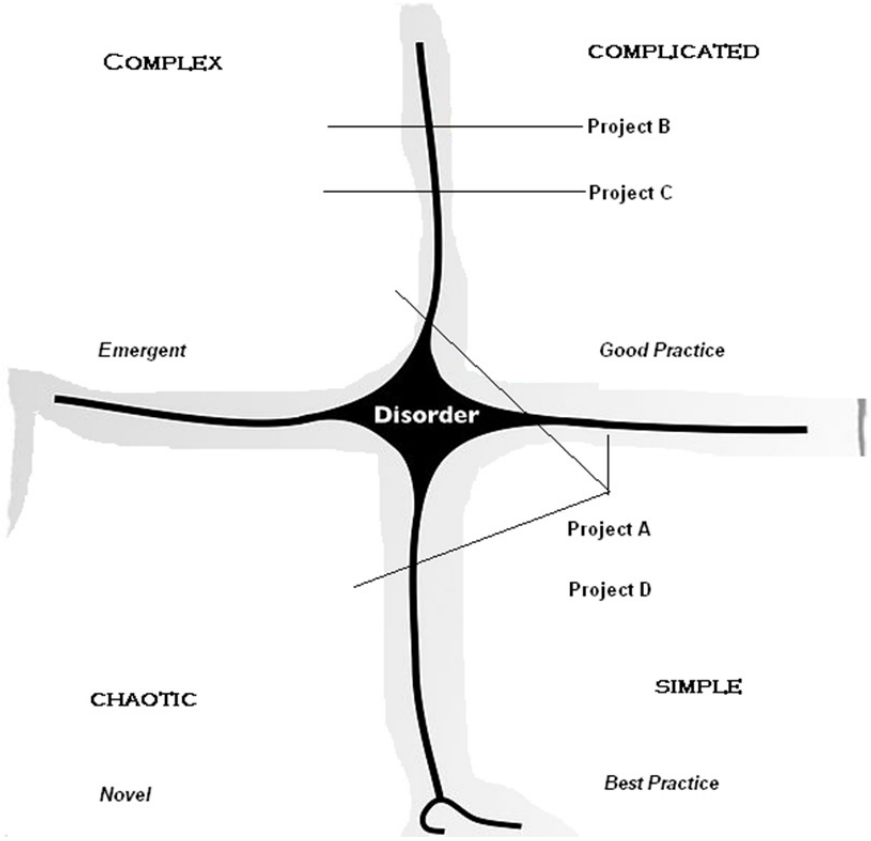
As the technological administration of SharePoint content types (that include metadata definition) and taxonomy controls were assigned to records managers the number of issues assigned to them was very high. As the user decision was needed on a number of issues many of them had to be allocated to user and records managers as well.
Other projects also required attention of records managers although users were not always aware of the role they had in what they initially considered a purely an IT project.

Retrospective application of Cynefin framework

Cynefin framework was developed by Dave Snowden as a sense-making model as oppose to a usual (matrix) categorization model (CognitiveEdge, 2011). It is a model that draws from knowledge management and complexity science (Childs and McLeod, 2013) and helps decision makers to make sense of a range of business problems and situations (Kurt and Snowden, 2003).
Child and McLeod (2013) as first in the field of records management use Cynefin to map people issues from their comprehensive AC+ erm project[2]. In doing this they find that "[the framework] provides a different lens through which to view, makes sense of and re-perceive the ERM challenge and offers a strategic approach to accelerating change". As Cynefin may be applied to analyse various situations it is used to retrospectively map all SharePoint implementation projects in Figure 1.

[2] http://www.northumbria.ac.uk/acerm (Access date: 30.08.2013)

Figure 1.



Project A as it was initiated as a simple records management improvement process may be placed in the simple domain as the solution was expected to be based on best practice as usually record management projects are. However when faced with the specific SharePoint environment the project issues resolutions had to be based on good practice where solution was good enough to work but far from perfect, for example enforcing column mandatory values through coding or even on emergent practice example of which is deriving a complex workflow for automatic filenaming. Finally close to the end of the project adoption issues emerged due to lack of time on most users' side which brought the project in chaotic domain. The situation was resolved by office administrator taking over all ERM responsibilities and performing records registration on behalf of all users. Such solution was never envisaged or planned, it was result of purely people factor – one’s attitude and self-initiative. The project has ended successfully with all stakeholders pleased with the final results and benefits.

Project B initially revolved around the creation of intranet presence however as the Intranet (SharePoint) working group was approving and overseeing the project records managers who are group's members guided the users on how to improve overall records management using the benefits of organizational metadata and taxonomy controls acquired during Project A. Thus it was expected that Project B will be either complicated or complex considering the large project scope and lessons learned during Project A. Project super-users were appropriately trained by both IT and RM staff and were given site owner permissions to manage their site independently.
Project C also started off as improvement of intranet presence however user recognized opportunity to improve both business processes and records management. As user possessed certain knowledge of programming he was able to run the project by himself requiring only limited IT and RM guidance and no direct assistance.
Finally Project D is placed in simple domain as its scope is limited to a fraction of office's records within single SharePoint library with no additional collaboration or any other requirements. It is assumed that this project will remain in simple domain as best practice acquired during Project A should cover the limited scope of Project D.

Conclusion

The main purpose of this brief case study was to show that using the Cynefin framework was possible in four different scenarios. Using such frameworks in projects may be very useful as they may provide direction even in most grave situations during an ERM system implementation. Furthermore it was demonstrated that people factor may be crucial for ultimate project solutions. Finally the underlying message is that records managers are in position to lead any ERM implementation, needless to say that it is their duty to greatly influence all ERM projects. Too often is their role undermined or non-existent while they may be in a perfect position to bridge the gap between users and IT in such projects by bringing in a strategic overview of information management needs and by gently relieving the overwhelming pressure on IT and ever so gently bringing user expectations to a reasonable level that would prevent unnecessary draining of organizational resources. Gaining project management skills or even software development skills (at least gaining insight into methodologies like Agile), applying frameworks and standards from other disciplines such as ITIL or perhaps rebranding themselves as recordkeeping informaticians (Upward et al, 2013) could surely strengthen the records managers role in organizations and ERM adoptions in particular.

References

Childs, Sue; McLeod, Julie. A strategic approach to making sense of the “wicked” problem of ERM. // *Records Management Journal*. volume 23 (2013), 2; pages 104-135

CognitiveEdge. The Cynefin Framework. 11.07.2011. http://www.youtube.com/watch?v=N7oz366X0-8 (Access date: 30.08.2013)

Kurtz, C.F.; Snowden, D.J. The new dynamics of strategy: sense-making in complex and complicated world. // *IBM Systems Journal*. Volume 42 (2003), 3; pages 462-483

Upward, Frank; Reed, Barbara; Oliver, Gillian; Evans, Joanne. Recordkeeping informatics: re-figuring a discipline in crisis with a single minded approach. // *Records Management Journal.* volume 23 (2013), 1; pages 37-50

Basic Components of Information System for a Contemporary Maritime Agent

Mila Nadrljanski
Faculty of Maritime Studies – Split
Zrinsko Frankopanska 37, Split, Croatia
milamika60@yahoo.com, mila@pfst.hr

Sanja Frkić
OŠ kraljice Jelene
Put mira bb, Solin, Croatia
sanjafrk@gmail.com

Gorana Tafro
Croatia osiguranje
Trg Hrvatske bratske zajednice 8, Split, Croatia
goranatafro@gmail.com

Summary

The Information system "shipAGENT" is conceptually a modular information system based on a client/server technology customized for user-friendly Windows. This modern system is suitable for use both in small shipping agencies with only one office and in large shipping agencies with remote offices and business units. The program is written in English and is modularly constructed using a relational database, leaving the possibility of using all modules or only some of them. At the user's request, the program could be supplemented with additional modules according to their specific needs.

Key words: shipping agent, information system, relational database.

Introduction

There is not much reference about a shipping agent in literature. Croatia is one of the few countries which have specific provisions on the conditions required to perform the activity of maritime agents in legislation[1] (Croatian Maritime Code, Articles 687-696). In most countries it is regulated by civil or commercial laws. In the Croatian legislation, the concept of a shipping agent is primarily regulated by the provisions of Maritime Code and subsidiary by the provisions of Civil Obligations Act and regulations of Commercial Law. Croatian law does

[1] The requirements for the conduct of a maritime agent are regulated in accordance with Articles 687-696 of Maritime Code.

not contain provisions which regulate the procedural position of a shipping agent; hence his procedural position is determined by jurisprudence. Jurisprudence confirms that the shipping agent is not passively legitimized for services that he ordered on behalf of and for the account of the shipping company. "Regarding international regulations, this book[2] covers primarily the UNCTAD's (The United Nations Conference on Trade and Development) minimum standards for shipping agents that are made by special intergovernmental assembly of the United Nations concerning the fight against maritime fraud. UNCTAD adopted these standards in March 1988, recommending their use. They are not obligatory but they can determine guidelines for state bodies. Croatia adopted these standards in Maritime Code and Regulations on Shipping Agents. Croatian Association of Maritime Agents brought them to their acts: Statute and Rules of Conduct." (Borčić, 1999)
Therefore, the verdicts from 1965, 1968 and 1969 issued by High Commercial Court of the Republic of Croatia are interesting from the point of taking such a legal position. A shipping agent[3] is actively legitimized which means that he is allowed to represent his principal in court, pursuant to special authorization[4]. The classification of shipping agents was induced by the fact that there are more tasks performed by shipping agents in maritime business.

The tasks performed by a shipping agent

Shipping agencies delivering goods to ports of call require a shipping agent at port to coordinate various details concerning international regulations on navigation. Therefore, agents should be familiar with all jurisdictions the shipper or co-signer visit in shipping; they should maintain certified paperwork and arrange all points of contact for the vessel.
Each port has unique service distributor’s procedures and it would be difficult to maintain port operations without the help of a shipping agent. Therefore shipping agents are engaged to provide on-site services based on knowledge of port operations. They can arrange pilotage, tug lines and schedules, as well as provide customs mediation. In short, shipping agent coordinates everything what is needed while the ship is in port. Ships enter and leave the port 24 hours a day and shipping agent works with all engaged parties to ensure the functioning on schedule.

[2] Borčić, Vojislav. Međunarodni i hrvatski propisi o pomorskom agentu. Rijeka 1999.

[3] Responsibilities of the shipping agent to the principal are: not presenting himself as an agent, not acting according to the instructions, non-performance, failing to notify the principal or keep his interest and avoiding responsibility.

[4] The maritime agent may sign a contract of employment of a ship for and on behalf of both contracting parties on the basis of an express authorisation of the contracting parties.

Planning

A shipping agent coordinates arrivals and departures of ships, suppliers' requests and synchronizes pilot's and master's directions. Maintaining a formal schedule is also agent's responsibility as well as providing of relevant information. He creates a crew list and a passenger list and coordinates the personnel required on a vessel for scheduled loading and unloading. He is a contact person for certified signatures.

Shipping operations

A shipping agent should be accustomed to all shipping regulations for all ports. He should ensure compliance with international shipping regulations, safety and policies of the ship's owner. The agent should be knowledgeable in all applicable publications which contain regulations and are certified by particular port of call authorities. He should keep copies of all documents, including records of complaints, insurance and inspection certificates.

Communication

A shipping agent is available during the hours slated for communication with the ship's owner. Besides being a contact for certified signatures he is also the main point of contact in port for all concerned, for the crew and port authorities. His communication duties include many relationships with various agencies in port, including business and public relations. All communications are recorded in a book for reference. It is possible to extend contact hours if required.[5]

The shipping agent must be familiar with the International Maritime Dangerous Goods Code, or IMDG[6], for the necessary retention and transportation of any hazardous materials (Croatian Maritime Code, Article 57).

Maritime agency services are services relating in particular to the navigation, ships and their employment with special regard to: the assistance and clearance of ships, mediation in concluding contracts of employment of ships, their sales and purchase, building and ship repair, insurance of ships and cargo, safeguard of the ships' interests, supply and manning of ships, care for crews and passengers (Croatian Maritime Code, Article 688).

Today's world displaces one paradigm by the other (a new one) and that is where change, development and differentiation happen. Today's companies need better and quicker ways of achieving results of quality managers (Nadrljanski, Batinica, 2010, p. 2).

Shipping agents should be versatile and have multitasking skills. In many ways, the skills possessed by managers are the most valued resources of the organiza-

[5] Modern electronic systems allow communication 24/7.

[6] IMDG Code is a set of rules for monitoring maritime traffic. It entered into force in accordance with Article 57 of Maritime Code.

tion. Lack of managerial skills can defeat the most successful activities and in many cases can lead to the collapse of the organization (Nadrljanski M., 2010, p. 23).

Shipping agent - definition and classification of rights and duties

Shipping agent is a very significant aspect of Maritime Code, and is a relatively new legal and economic category (Contracts of Maritime Agencies, Articles 687-696). There is a common saying that the importance of something is best explained by the history of its development and the struggle for survival (Katičić, 1951). Throughout history, duties of a contemporary shipping agent were performed by a ship owner who was also the commander of a ship. Through navigation development, the functions of ship owners and masters have been separated. Ship owner starts managing navigation from the land, while master takes care of all the administrative formalities when entering or leaving port, procurement of cargo to be transported, and its loading and unloading. With further development of shipping and trade, these duties were assigned to a shipping agent[7]. In navigational operations agents perform under the general or special authorization in the name and for the account of the principal, and not on their own behalf and for their own account (Skorupan, 2006).

Legal Terms

In contemporary commercial and maritime practice, the contracting shipper never performs all obligations under the shipping contract personally. Most shippers are corporations which act only through their agents. It is essential that all contemporary shippers subcontract with other companies in order to perform specific aspects of shipping business or its part. Shipping agent is a legal or natural person registered in Shipping Agent Registry to conduct shipping agency activities, in accordance with the provisions of Maritime Code[8] (Ordinance on Shipping Agent's Performance, Rights and Duties, Article 5).

Employee agent is a natural person holding the license of a shipping agent who on behalf and for the account of a shipping agent directly performs shipping

[7] Duties regulated in Article 689, Croatian Maritime Code: Under a maritime agency agreement based on a general authorisation the maritime agent undertakes, for and on behalf of the principal, to perform all services or all services of a specific type in the field of his activity. A maritime agency agreement based on a general authorisation must be drawn in writing. An agreement that is not drawn up in writing shall have no legal effects. In case of doubt as to the limits of the general authorisation it shall be deemed that it refers to the clearance activities. If a maritime agent deals exclusively with mediation or with the stipulation of contracts of employment of ships, in case of doubt as to the limits of the general authorisation, it shall be deemed that it refers to the mediation for stipulating such agreements, excepting bare boat charters, charters by demise and time charter for a whole ship.

[8] Maritime Code (Official Gazette 181/2004) was adopted on 8 December 2004, and entered into force on 29 December 2004.

agency activities, particularly dispatching activities. To meet requirements of the Ordinance, he has to dispose of his own funds of at least 120,000 HRK, or other adequate guarantees in banks, financial institutions or other companies.
These guarantees may be replaced by the appropriate national or international insurance. Guarantees or insurance must be available for the entire duration of his activity as a shipping agent.
Maritime Code defines contract of maritime agency as a contract by which the maritime agent undertakes to perform maritime agency services and services of assistance, mediation and representation, for and on behalf of the principal, on the basis of a general or special authorization, whereas the principal undertakes to reimburse the maritime agent for the expenses and pay to him a remuneration. It is specified which services can be classified into maritime agency services. Maritime agents perform a very wide range of tasks, or may participate in mediation in concluding majority of transactions related to ships, and they have a significant role in assisting the commander of the ship, crew and passengers. He is a person who represents or intermediates in maritime business based on maritime agency agreement. The shipping agent is an institution within maritime law and maritime economy, as it is stated in literature (Borčić, 1992).
A shipping agent has many duties:

- handling ship formalities during the arrival and departure (he is among the first to arrive on the ship)
- collecting data on cargo, crew and port requirements
- procuring supplies, crew
- taking care of medical services and other needs of the crew
- supervising cargo
- completing documentation upon leaving
- taking care of the bill of lading, accounts, freight, customs and other documents
- ordering towing and piloting
- arranging ship repairs

Information system used by a shipping agent - agency

Successful operation of a maritime company and successful managing of a shipping agency are inconceivable without adequate data, information and knowledge. Therefore, data, information and knowledge can be comprehended as a kind of resource of a maritime agent.
Interest in the problems of managing information resources is increasing rapidly in recent years. Until recently, maritime agents dealing with problems of data management used to have only a passive role at the level of the relationship with their clients. However, it was realized that we should develop new operating relationships between information system used by an agent and other parts

of maritime company - the relationships that will be a part of the company organization, which in turn, would improve achieving its objectives.
Data and information are perceived as organizational resources that should be invested in, and which should be taken care of. The mentioned changes and improvements are gradual and require carefully planned and professionally designed programs, which depend on the shipping agent's intended use and control of information.
Strategic planning of information resources is used for:

1. Linking information technology and information systems with strategic business planning of maritime economy;
2. Assisting in creation of control mechanisms for plan implementation process;
3. The creation of an information framework which will provide further analysis and design. This will allow coordination of specially developed database and information systems and ensure that the information they provide is consistent.

A shipping agent has to act expeditiously in making decisions. The results of his actions could be disastrous or, on the other hand, lucrative for the financial status of his principal. In order to be updated, he must have a constant and fast access to any information needed. Using shipping software improves quality and saves time. It can help a shipping agent to prepare and organize ship and port documentation, generate reports, track company's business in history and use communication services, such as an e-mail, fax, telex and video communication in a manner most suitable to his work requirements.

The shipAGENT PRO system

For this purpose there are programs used to automate maritime agent's tasks. In this paper we are going to illustrate an information system called the shipAGENT[9] (http://www.edss.ee/ShipAgentPro, access 2012). It is a modular information system developed by using modern tools for the development of information systems based on client / server technology and is performed using user-friendly Windows. The system is suitable for use both in a single agency office as well as in large shipping agencies with remote offices and business units. The program was created in Croatia, but it is written in English language. Since it is created modularly, using a relational database, it is possible to use all or only some of the modules. On clients' requests, the program can be supplemented with additional modules according to their specific needs. ShipAGENT system is communication-oriented and it can be installed in one or more work-

[9] The ShipAGENT PRO system is a software package developed by InfoExpert d.o.o. 51000 Rijeka, Lorenzov prolaz 13. (16 August 2011)

stations (PCs) via a local area network (LAN). In remote offices / business units it uses modem connection and telephone lines.

Modules in the shipAGENT PRO system

- Call – This is the basic module of the program that can be used independently. It consists of all data related to the ship's visit in a port of call from its notice of arrival to its departure. It contains a lot of data concerning the ship itself, entering and leaving the port, Statement of Facts, the quantity and type of the loaded and unloaded cargo, expenses and invoices, manifest, bills of lading, and so on. The module can be used independently but its true value is in combination with other modules.
- Cargo – This is a module for monitoring announced and performed loading and unloading. It includes the data of cargo operations, cargo types, cargo units, quantity, cargo codes and cargo description. The number of operations and cargo types per call is not limited.
- Deposit – It monitors information about requested and security deposit: payer, the date of the requested deposit, the amount requested, the currency, the amount paid and method of payment, bank statement number and the date.
- Facts – It is a module that tracks all data related to significant events on board and in port, necessary for the document Statement of Facts. It includes: the date and time of inception, the date and time of completion, the code and a description of events.
- Invoices – This module tracks all expenses and bills of the ship, including agent's expenses which are transferred into Disbursement Account. The data include: service provider, payer of services, types of services, the number and date of the document, quantity, the price in USD and HRK, Value added tax (VAT) and reclamation label.
- Extra Invoice and Credit Note – It collects data and prints additional accounts and various Credit Notes.
- Manifest – This module is used for collecting all the data for the preparation and printing of cargo and freight manifests.
- Bill of Lading (B / L) – It collects all the data for the preparation and printing of standard bill of lading.
- Exchange Rate – It collects information about buying rate, middle rate and selling rate for all major currencies.

The ship AGENT program applies a large number of different codebooks. The most important are:

- Party – This is a separate database, integrated in the shipAGENT program, and includes data of all business partners doing business with an agent. All of them are labelled with business role, meaning that one partner can have multiple roles (owner, principal, shipper, receiver, for-

warder, agent, service provider, P & I club, surveyor). It contains: partner's code, partner's name, country, address, postcode, city, Post Office Box; Telephone: country code and city code, phone number 1 and 2, fax number 1 and 2, mobile phone number, e-mail address, contact person's name and identification number.

- Ship – This is a separate database, integrated in the shipAGENT program with the data of all the ships agent works with. It contains: ship's code, ship's name, flag, the owner, the port of registry, Call Sign, GT[10], NT[11], DWT[12], LOA[13], Draft, Type, Built, Capacity (Grain, Bale, TEU[14]) and IMO No[15].
- Other codebooks: Flag - according to ISO codebooks[16] of countries, Port - according to the UN codebook of ports[17], Cargo Type, Cargo, Ship Type, Berth, Service type, Currency – according to ISO letter and numerical currency code list, Call type, Fact, Branch Office, Clause, Unit, Packing and other additional internal codebooks. They can be created and edited on request.

Program functionality

Data Entry

Holding the cursor over the fields while entering data, you get a short instruction (mini help) on how to enter data.

Browsing the database

There is a possibility of looking for your business partner by his name (or part of it), country, type of partnership and the city.

Ships can be found by their name (or part of it), flag, their type and the owner.

[10] Gross tonnage – an index related to a ship's overall internal volume and calculated by measuring a ship's volume from keel to funnel, to the outside of the hull framing.

[11] Net tonnage

[12] Deadweight tonnage – the amount of weight a ship can carry without riding dangerously low in the water.

[13] Length overall – the maximum length of a vessel's hull measured parallel to the waterline.

[14] Twenty-foot equivalent unit – a unit of cargo capacity often used to describe the capacity of a standardized intermodal container.

[15] International Maritime Organization numbers – unique identifiers for ships, for registered ship owners and management companies.

[16] International Organization for Standardization

[17] United Nations Code for Trade and Transport Locations (UN/LOCODE); http://www.unece.org/cefact/locode/service/location.html (access 2013)

Ports of call can be browsed by year, the business unit, dispatch number, the ship, and the principal.

Reports and documents
There are various lists which could be obtained from the database (for business partners, ships ...): conditions and movement of ships, invoice, manifest, bill of lading, statement of facts, deposits, financial statements (by party, debtors, fee report - with criteria, fee totals - cumulative or with details), statistics and graphs, fax and e-mail reports.

The program is functional and widely applicable. The configuration necessary to work with the shipAGENT is not demanding: personal computer (PC) with 16 MB RAM (recommended Pentium processor), Windows 95 or Windows NT operating system, modem for remote maintenance, connection to the local area network (for multiple units) and a printer. It is easy to learn how to use a program for a person who has basic computer skills and is familiar with Windows and the Internet.
Although the program is useful and greatly facilitates shipping agent's tasks, his managerial abilities in overall business process should not be neglected. The main objective of managing a business process is not automation, but focus on the processes themselves and effective management. It does not necessarily imply automation but optimization. Business processes are a key factor in the integration of the company organization, while the information technology is just supporting tool in business processes (Nadrljanski Đ, Nadrljanski M, 2010).

Conclusion

Duties and responsibilities of a shipping agent are regulated either by Maritime Code (in Croatia) or by civil and commercial laws, as it is elsewhere in the world. This fact emphasizes severity and complexity of tasks performed by a shipping agent. Agents should be versatile and have multitasking skills. In order to simplify and organize such demanding tasks, there are computer programs which can be used by single agency offices as well as large shipping agencies with remote offices and business units. The ShipAGENT is created modularly, using a relational database, so it is possible to use all or only some of the modules. It can be supplemented with additional modules according to specific needs of a client. It is functional, easy to use and does not require complex operating system.

References

Borčić, Vojislav. Međunarodni i hrvatski propisi o pomorskom agentu. Rijeka: Udruga pomorskih agenata Hrvatske, 1999

Borčić, Vojislav. Pomorski agent u zakonodavstvu Hrvatske. // *Zbornik Pravnog fakulteta Sveučilišta u Rijeci*. 13 (1992); p. 13-23.

InfoExpert d.o.o. The Ship Agent PRO system. Rijeka, 2011, http://www.edss.ee/ShipAgentPro (Access 2012)

Katičić, Natko. Ugovori o iskorišćivanju brodova na moru, pravni značaj i podjela. // *Zbornik rasprava, Jadranski institut JAZU*, Zagreb, (1951), p.14

Nadrljanski, Đorđe; Nadrljanski, Mila. Informacijsko komunikacijski sistemi i tehnologije. Split: Redak, 2012; p.9

Nadrljanski, Mila. Komunikologija i menadžment. Split: Redak, 2010; p.23

Nadrljanski, Mila; Batinica, Vicko. Education of Maritime Managers. Opatija: *MIPRO, 2010 Proceedings of the 33rd International Convention,* pp 954 – 958 (2)

Pomorski zakonik RH = Croatian Maritime Code. Articles 57, 688, 689, 696

Pravilnik o uvjetima za obavljanje djelatnosti pomorskog agenta, te pravima i obvezama pomorskog agenta=Ordinance on Shipping Agent’s Performance, Rights and Duties. Article 5

Skorupan, Vesna. Odluke domaćih sudova: Naknada za korištenje objekata sigurnosti plovidbe // PPP god.45 (2006), Vrhovni sud Republike Hrvatske Presuda broj: II Rev-123/1995-2. Articles: 160, 162-164

Ugovor o pomorskoj agenciji = Contracts of Maritime Agencies (Chapter III of Croatian Maritime Code). Articles 687-696

United Nations Code for Trade and Transport Locations (UN/LOCODE); http://www.unece.org/cefact/locode/service/location.html (Access date 20 June 2013)

Information Technology as Basis for the Changes in the Social Role of Film

Marcella Jelić
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
marcella.jelic@gmail.com

Summary

Digital technologies are changing the way we see film as a mass medium turning it into one of the most important media of today, and, judging by the current trends, of tomorrow as well. Mostly made with digital equipment, the largest number of films that appear today are short, often non-narrative, even more often amateur. Paradoxically, it seems that the proclaimed 'death of film tape' could mark a new, more active 'life' for the film.

Key words: social role of film, film as mass media, digital technology

Introduction

Digital technologies are changing the way we see film as a mass medium. According to some theoreticians, with the invention of radio, and television especially, the film has lost its function of a medium of mass communication. But is it really the case? Can we now, in the era of massive use of digital technologies and the expansion of new media declare the film to be forever lost to the mass media? We will try to prove the opposite: that the new technologies have turned the film into one of the most important media of today, and, judging by the current trends, of tomorrow as well. We will try to prove that film is one of the means of mass communication, and if the existing trends of increasing dominance of audio-visual content continue to develop, the film, or "cinematic way of seeing the world" could take more and more important role in the future. We will try to prove that, along the lines of changing technology, the social role of individual media is also changing, and the theory of mass communication should constantly review the new circumstances. Our aim is to investigate the reasons film is such an interesting medium today.

Changes in the social role of film

"Times have changed, but in a world that has become *hypermediatized* the social role of film, contrary to what is sometimes claimed, is not by any chance on a descending path. Nowadays we actually turn to film when we are trying to awaken the conscience and to measure big institutions. In this way the film helps international organizations, as is the case with the International Human

Rights Film Festival, which was explicitly organized in response to the ineffectiveness of the Council of the United Nations in that domain."[1] Although it was created a little more than 100 years ago and has been continually evolving to this day (silent, sound, color, 3D), the social role of film has already transformed considerably, to the extent that some theoreticians believe that we can no longer perceive film as a medium of mass communication, but only as a work of art.[2]

Given the speed of its development, it took quite some time for film to become accepted as an art. In words of Rudolf Arnheim, film was animated image, somewhere between theatre and still photography.[3] Domenico Tumiati[4] argued that the future of film would not be artistic, because art is individuality, and film is a machine. At first, the film was perceived solely as entertainment for the (ignorant) masses, then it became important means of delivering information – not just about the world news, different countries and cultures, but also about the lifestyle and social values of other nations, easily accessible and understandable to everyone, only to emerge at last as an important art. The term "the seventh art" was coined in the year 1911 by one of the first film theoreticians and critics, Ricciotto Canudo, who put film on equal footing with the other arts – poetry, architecture, music, dance, painting and sculpture.[5]

There are film theoreticians who claim that social role of film has changed and that it ceased to be a medium of mass communication, and Hrvoje Turković is among them. "Although there are people old enough to remember it, it is quite difficult to grasp that film once performed almost all programming functions of television. The reason is the slow but inevitable evolution of film, turning into a media system identical to other 'art' systems (literature, visual art, music and theatre). What has quietly but thoroughly changed was the change of the civilizing role of film considered in its totality. What was the system the film abandoned? It was the system of the so called media of mass communication (also called the mass media –today paradigmatically represented by television, radio and newspapers),"[6] says Turković and recalls that film formed as a mass medium at the time of cheap cinema, when the picture show consisted of program of short films, which usually contained current information (newsreels, reports), as well as commercials, advertisement and educational films. "During this period, the audience developed that habitually, regularly (daily, weekly) went to the cinema, following the key changes of programmes (sequels, new editions of

[1] Lipovetsky, Gilles, Serroy, Jean, Globalni ekran, Novi Sad, Akademska knjiga, 2013., 333.

[2] Turković Hrvoje, Narav televizije, ogledi, Meandar, Zagreb, 2008.

[3] Arnheim, Rudolf, Film i stvarnost, Dušan, Stojanović ed., Teorija filma, Beograd, Nolit, 1978., 111.

[4] Stojanović, Dušan ur, Teorija filma, Beograd, Nolit, 1978., 13

[5] Ricciotto, Canudo, Estetika filma, Dušan Stojanović ed, Teorija filma, Beograd, Nolit, 1978., 54.

[6] Turković, Hrvoje, Narav televizije, ogledi, Zagreb, Meandar, 2008., 175.

journals, various 'innovations' in feature films etc.). In short, film had the exact same function that television and radio have today and the newspapers have had from the beginning. And why has film lost that function? There are many reasons, but two are, it seems, the most prominent. The first reason is the time which elapses from recording to screening, and the other is that showing and viewing the film was an inconvenient act" states Turković[7], adding that the film of the 'non mass' period has not ceased to be massively popular, but only ceased to be a means of mass communication as it was at the beginning.[8] "Let me repeat: it is now marked by fragmentation and individuality (disciplinary, authorial, stylistic), reduced relevance and permanent availability (movies are watched again after being shown in the cinema, on television, video, for teaching purposes, for the purpose of theoretical analysis) and is characterized by certain optional nature, selective reception (we go to the cinema and watch movies occasionally, not because we have an inherent social obligation to do so). The social role of the film has changed," Turković says.[9] But after the change diagnosed by Turković in the early 90s,[10] today in the time of expansion of digital technologies and the impact of new media, which leads to the formation of hybrid media, the social role of film is changing all over again.

Film as a medium of mass communication

According to the definition from the Film Lexicon, the film is, among other things, "a comprehensive and socio-experiential product".[11] From the point of view of theory of mass communication, the above-mentioned social function of the film is particularly important. In fact, some film theories[12] and many theories of mass communication[13] recognize communication as one of the functions of film. "...film screening, like every other screening as a social phenomenon, belongs to a special category of social phenomena – the phenomenon of communication".[14]

Mass communication is one type of communication (in the most common divisions the remaining types are interpersonal communication and intrapersonal

[7] Ibid, 177.

[8] Ibid, 181.

[9] Ibid, 182.

[10] Altought the book Narav televizije (Turković Hrvoje, Zagreb, Meandar) was published in 2008. the text was written in 1992.

[11] Filmski leksikon, Leksikografski zavod Miroslav Krleža, Zagreb, 2003., 177.

[12] Turković Hrvoje, Teorija filma, Meandar, Zagreb, 1994.

[13] McQuail Denis, Mass Communication Theory, 6th edition, Sage, 2010. and Dominick Joseph R. The Dynamics Of Mass Communication, 2nd edition, Random House, 1987.

[14] Turković, Hrvoje, Teorija filma, Zagreb, Meandar, 1994.,59.

communication[15]) and it is defined as "the process by which a complex organization with the aid of one or more machines produces and transmits public messages directed toward large, heterogeneous, and scattered audiences".[16] The mass media are indispensable as a channel that transmits messages in the system of mass communication. In addition, film has all the properties of a "mass communicator ": mass communication is produced by a complex and formal organization, there is a number of "gatekeepers" (a person or group of people responsible for the selection of content that will reach the audience), it needs significant financial support to function, it exists in order to generate profit and it is competitive.[17] One does not necessarily has to agree with the last statement, because there are mass media the purpose of which is not to make profit. "The media can be local, national and international, private, public, governmental or community, i.e. associative or third sector media as well as mainstream and alternative ".[18]The film meets this definition as well, since numerous works of film art, experimental film especially, are made without any ambition to be profitable. In fact, almost every film, some rare exceptions excluded, is produced by a production company or, in Hollywood, by an influential film studio owned by a multinational corporation (complex and formal organization) , the "gatekeepers" are film agents, producers and directors who, based on various criteria, select the projects to be implemented, and even the cheapest films are relatively expensive (among the cheapest is the Croatian film *Show Must Go On* by a young director Nevio Marasović produced with a modest amount of about 200,000 kuna[19]). Production companies are usually privately owned companies and their primary (though not necessarily the only) interest is profit-making, and it is especially evident as far as competitiveness is concerned, not only in the struggle for audiences in theatres, but in the fight for public funds that are financing film projects (in Croatia, this fund is Croatian Audiovisual Center).To sum up, from the theoretical point of view, film can be viewed as a medium of mass communication.

Information technology and film

Even if we agree with the statement that film has lost its role as a mass medium due to its time delay and the impracticality of the act of screening (going to the

[15] Brittner John R., Mass Communication An Introduction, 5th edition, Prentice Hall, Englewood Cliffs, 1989., 10.

[16] Dominick Joseph R., The Dynamics Of Mass Communication, 2nd edition, Random House, 1987., 16.

[17] Ibid, 20.

[18] Preuško, Zrinjka, ed., Uvod u medije, Zagreb, Jesenski i Turk, 2011, 15.

[19] Saračević, Igor, Pulu otvara potpuno drugačiji hrvatski film, Tportal.hr, 7th, July, 2010. http://www.tportal.hr/showtime/film/75787/Pulu-otvara-potpuno-drugaciji-hrvatski-film.html (10th July 2013.)

cinema), it should be noted that Turković published this in 1992, the same year in which Croatia became a part of the global computer network, the Internet, when international communication connection which connected CARNet Internet exchange point in Zagreb to Austria was established.[20] So the statement might have been correct at the time – new digital technologies and, above all, massive use of the Internet has changed the role of film once again. "The tradition of the printed word, initially prevalent in the language of cultural interfaces[21] is becoming less important, while the part played by cinematic elements is progressively getting stronger. This is consistent with a general trend in modern society towards presenting more and more information in the form of time-based audio-visual moving image sequences, rather than in textual form," states Manovich[22], noting that it is thanks to computer that film became a visual Esperanto.

Manovich takes one step further, noting that the film aesthetic strategies have become fundamental principles of computer software, while Gilles and Jean Lipovtsky Serroy note that in an era of omnipresent global screens – screens in shops, on airports, in restaurants, coffee shops, cars, airplanes; all screen sizes, mobile screen; the screen on you, the screen with you; a screen for all purposes, the screen to see everything, video screen, miniature screen, graphical screen, mobile screen, touchscreen[23] – "the man of hypermodern society sees the world as a film; to him, the film provides glasses which he unconsciously uses to look at reality in which he lives. The film has become the thing that shapes the global view of various spheres of contemporary life".[24]

The appearance of digital technology, computers, other screens, and, even more importantly, the Internet, has transformed the existing mass media and opened up space for emergence of hybrids and new media (web portals). As Roger Fidler writes, thanks to the technology, first and foremost, "mediamorphosis" happens – "transformation of communication media usually occurs due to complex interweaving of the expressed needs, competitive and political pressures and social and technological innovations".[25] Convergence and multimediality have paved the way for audio-visual information, business and leisure activities on the new platform, the Internet, which has become available to a substantial part of the world's population, not only on computers but also on mobile phones

[20] http://www.carnet.hr/o_carnetu/o_nama/povijest_carneta (1st July, 2013.)

[21] According to Manovich cultural interfaces are ways in which computers present cultural dana and enable interaction with it. Manovich, Lev, Language of New Media, Cambridge, The MIT Press, 2001., 70.)

[22] Manovich, Lev, Language of New Media, Cambridge, The MIT Press, 2001., 78.

[23] Lipovetsky, Gilles, Serroy, Jean, Globalni ekran, Novi Sad, Akademska knjiga, 2013., 24.

[24] Ibid., 42.

[25] Fidler, Roger, Mediamorphosis, Beograd, Clio, 2004., 41.

and tablet devices. We are connected through the Internet constantly, at any place and any time, and the content consumed online, especially social networks and web portals, are increasingly focused on short audio-visual forms – segments of TV shows, music videos, commercials, advertising films, short video clips recorded either by professional directors, cameramen, journalists or by amateurs, readers who send their work to the newsroom or upload it themselves to social networks and video-sharing websites. It is mostly short and informative, entertaining or advertising film material, essentially very similar to materials shown in theatres in the early days of film. Oftentimes it is a short recording of "real life", which basically does not differ much from the first movies made by Lumiere brothers, for example. "Members of the Lumiere family patented portable, we would say – compatible portable camera – already in the year 1895, as amateur photographers/cameramen, recording all that was happening in the private and public life, and they also 'provoked' portraits of close friends and relatives. They reproduced life...".[26] However, what "Workers Leaving the Lumiere Factory " once was, today is a funny amateur footage of children playing, such as "Charlie bit my finger again"[27] – one of the most popular video clips on YouTube with more than half a million views[28], which in aesthetic terms has much in common with the early film recordings – static camera, short form, the entire film consisting of one take and documenting a piece of "reality".
Thanks to the sites like YouTube and Vimeo (which are not necessarily considered to be mass media because the materials published there mostly belong to private persons, not "complex organizations", as required by the mass media definition), whose video materials are integrated into portals, web editions of daily newspapers, magazines and television programs, the private, amateur footage became a part of the "official" media. In short, through their integration into "classical" mass media, the video clips from social networks have become institutionalized. Editors of mass media are selecting and integrating them into their web portals and magazines, eliminating a possible dispute over whether such audio-visual material could be considered a part of the media of mass communication. It certainly could, but this phenomenon has a dark side. "With its vanity and absurdity, the YouTube content has managed to overshadow even blogs. Nothing is too prosaic or narcissistic for those monkey-videographers. The site is an infinite gallery of amateur films that present the poor fools dancing, singing, eating, washing up, buying, driving, cleaning, sleeping or just staring at their computers,"[29] remarked Andrew Keen, highlighting, among other things, the problem of relationship between professionalism and amateurism that will prove to be the key to understanding the development of the media.

[26] Miltojević, Branislav, Podeljen ekran, Beograd, Filmski centar Srbije, 2011., 52.

[27] http://www.youtube.com/watch?v=_OBlgSz8sSM (1st July, 2013.)

[28] http://www.youtube.com/charts/videos_views?gl=US&t=a (1st July, 2013.)

[29] Keen, Andrew, Kult amatera, Zaprešić, Fraktura, 2010., 21.

While the average quality of works is up for discussion, the hyper production has had at least one positive effect: it has contributed to the return of the film as one of the most important media of today. "*Omniscreen* is not the films' tomb: more than ever, the film proves its imagination, diversity, vitality".[30] On YouTube, 10 million videos are viewed daily and 65,000 new ones are added, "never so many recorded sequences have been produced and displayed, never so many artistic and authorial videos have been made, never the audience have become global so fast".[31] The short cinematic form that has had a special significance in the formation of new role of the film as one of the dominant media is certainly the music video. Lev Manovich notes the emergence of new forms of film that developed in the 80s: they are non- narrative, shown on television or computer screen rather than in a movie theatre, and at the same time they are denying the film realism.[32] He mainly has music videos in mind, because they have "since the eighties brought a new way of looking at things to film, a new and completely different way of showing and narrating".[33] As far as music videos are concerned, a stylistic turning point was the music video for Michael Jackson's *Thriller*, directed by John Landis in 1983. With its enormous half a million dollar budget and integral 14-minute-long version *Thriller* "opened a new chapter in the history of music industry, establishing the concept of music videos shot in the form of short film".[34] Meanwhile, video clips have moved online in significant numbers, to YouTube, Vevo and other video-sharing websites which have, with their accessibility on multiple platforms, substituted the already weakened role of music television. Even the popular MTV stopped using the slogan "Music Television" in 2010. At the same time, the clips watched on YouTube most are music videos; nine out of ten most popular videos are music clips,[35] which is the content lacking on Croatian national TV stations (CMC excluded). The majority of the music videos today are produced almost exclusively for the Internet broadcast and they¸ are typical hybrid media somewhere between television, film, music and the Internet.

"Art (computer art), music (music video), game (videogame), commercials, chat, photos, knowledge – nothing can escape digitalized networks of the new *screenocracy*. Out entire lives, all our relations with the world and with others are more and more mediated by numerous interconnections through which screens never stop to converge, interact and link mutually".[36]

[30] Lipovetsky, Gilles, Serroy, Jean, Globalni ekran, Novi Sad, Akademska knjiga, 2013., 26.

[31] Ibid, 32.

[32] Manovich, Lev, Language of New Media, Cambridge, The MIT Press, 2001., 310.

[33] Lipovetsky, Gilles, Serroy, Jean, Globalni ekran, Novi Sad, Akademska knjiga, 2013., 306.

[34] Miltojević, Branislav, Podeljen ekran, Beograd, Filmski centar Srbije, 2011., 99.

[35] http://www.youtube.com/charts/videos_views?gl=US&t=a (1st July, 2013.)

[36] Lipovetsky, Gilles, Serroy, Jean, Globalni ekran, Novi Sad, Akademska knjiga, 2013., 36.

Today, at times of constant exposure to various and omnipresent screens and constant online availability, only expected to further increase in the future, there are no more obstacles to practicality of watching a movie: it can now be watched anytime and anywhere, alone or in the company. Online, downloaded, streamed, and at all screens available: in theatres, on television, computer, tablet, mobile phone. "The smaller the audience visiting darkened auditoriums, the greater is the desire to record, there is more cinematic narcissism, but also the greater are the expectations of the visual, of hypervisualisation of the world and of oneself. We are not satisfied with watching the "big" movies any more, but we want to watch movies about the moments of our lives and what is happening right now. It is not a denial of film, but the expansion of the film spirit in the globalized *film-vision*. Omniscreen is not degrading the film, on the contrary: it contributes to the expansion of the film perspective, doubling the life of the moving image, creating a general and widespread *film-mania*".[37]

Amateurism vs. professionalism

One of the most important phenomena of strengthening the new media and digital technology proved to be the so-called democratization of the media and the strengthening of amateurism in various fields: in journalism,[38] film, music, software production. Andrew Keen warns that any phenomenon in which the audience has the final word is a very dangerous and harmful one,[39] and the victims are not only professionals, because amateurs "diminish the value and take away their jobs", but all users of this "'free' content constantly reaching for our attention".[40] He argues that information produced by amateurs is unreliable, unverified and that the democratized media are eventually forcing us to become critics and amateur editors, and regard everything we read with a dose of scepticism; we pay for "free" information with our most valuable resource – our time, Keen argues.[41] His opinion is supported by Nicholas Carr, who explains the genesis of the “cult of the amateur” as a result of the combination of cheap and easy-to-operate digital equipment and infinite storage capacity of data in digital form. "The major constraints in supply of creative works – high costs and narrow distribution channels – are disappearing. Because the most common cultural goods consist of words, images or sounds, all of which can be expressed in digital form, they are becoming as cheap to reproduce and distribute as any other in-

[37] Ibid, 39.

[38] Amateurism is very important in film criticism. Today 'collective grade' on Internet Movie Database is more important than opinion of professional film critic. "Millions of ordinary people are the new creators of taste" wrote Chris Anderson (Anderson, Chris, Dugi rep, Zagreb, Jesenski i Turk, 2008., 120.)

[39] Keen, Andrew, Kult amatera, Zaprešić, Fraktura, 2010., 39.

[40] Ibid, 61.

[41] Ibid

formation product. Many of them are also becoming easier to create, thanks to the software and online data storages and inexpensive production tools such as camcorders, microphones, digital cameras and scanners. Tasks that once required a lot of money and training, from film developing through video editing and graphic design to sound mixing can now be performed by amateurs in their dens, offices and schoolrooms. The proliferation of blogs, podcasts, video clips and mp3 files testifies to the new economics of cultural creation. And all the new digital products, whether fashioned by professionals or amateurs, can find their place in the online store. The virtual shelves of the Internet can expand to accommodate just about anything".[42]

Understanding the new relationship between the amateurism and professionalism could be the key to understanding the future transformation of the mass media because, as explained by Gilles Lipovetsky and Jean Serroy, proliferation of media offers and expansion of computerized communication change the fact that more individuals have access to the media in hyperindividualistic way following their desires, moods and their individual schedules.[43] "Global screen is perceived as an instrument adapted to the special needs of each individual: communication one-to-everyone is followed by all-to-all, mass-media is followed by *self-media*".[44] We should not be surprised, says Chris Anderson, if some of the most creative and influential works in the next few decades emerge from inspired hobbyists, rather than from traditional commercial sources.[45]

There is no single answer to whether democratization of media does more good or harm, but it is certainly changing the paradigm. Perhaps the time has come to accept the changes and modify the definition of mass media.

Conclusion

Digital technology and the Internet have brought about yet another change in the way we understand film. The existing tendency of visual (over)saturation leads us to conclusion that film is an omnipresent medium, and therefore, due to its persuasiveness, perhaps the most influential. In times when the interest in print media is declining, the radio has lost its influence and the television, while still very relevant, is undergoing an advanced stage of digital convergence, and the same goes for printed media, publishing, music and news agencies, the Internet has, as we have already stated, played a crucial role in the affirmation of film as one of the dominant mass medium of today. It was the Internet, combined with digitalization, that has enabled the omnipresence of film content.

[42] Carr Nicholas, The Big Switch: Rewiring the World, from Edison to Google, New York, W.W. Norton & Company, 2009. str. 150.

[43] Lipovetsky, Gilles, Serroy, Jean, Globalni ekran, Novi Sad, Akademska knjiga, 2013., 287.

[44] Ibid, 288.

[45] Anderson, Chris, Dugi rep, Zagreb, Jesenski i Turk, 2008., 76.

It is important to point out that what we consider to be a film is, in fact, a "total and interconnected combination of characteristics of all film works, film footages and film images of the whole so-called film corpus", states one of the definitions in the Film Lexicon. Most of the films today are made with digital equipment, so it should be kept in mind that perhaps the largest number of films that appear today are short, often non-narrative, even more often amateur... Paradoxically, it seems that the proclaimed 'death of film tape' could mark a new, more active 'life' for the film.

Based on everything stated, film can be considered a mass medium. Perhaps a part of its power lies in the fact that since the very beginning the film has never discriminated against amateurism, while cinema clubs have played an important role in spreading the popularity of film and the preservation of film art. Thanks to accessible and cheap digital equipment films are recorded easily and quickly with cameras or mobile phones, edited on computers, distributed online, and since there is no discrimination of professionals against amateurs, everyone feels invited to participate in the production. There is no doubt that development of new technologies is what the future holds, together with improvement of the existing ones. The aim is, at least when film is in question, to achieve even more realistic representation (perhaps even without the mediation of the screen) and to come up with new profit-making models.

References

Anderson, Chris, Dugi rep, Zagreb, Jesenski i Turk, 2008.

Brittner, John R., Mass Communication An Introduction, 5th edition, Englewood Cliffs, Prentice Hall, 1989.

Carr, Nicholas, The Big Switch: Rewiring the World , from Edison to Google, New York, W.W. Norton & Company, 2009.

Dominick, Joseph R., The Dynamics Of Mass Communication, 2nd edition, New York, Random House, 1987.

Fidler, Roger, Mediamorphosis, Beograd, Clio, 2004.

Filmski leksikon, Zagreb, Leksikografski zavod Miroslav Krleža, 2003.

Keen, Andrew, Kult amatera, Zaprešić, Fraktura, 2010.

Lipovetsky, Gilles, Serroy, Jean, Globalni ekran, Novi Sad, Akademska knjiga, 2013.

Manovich, Lev, Language of New Media, Cambridge, The MIT Press, 2001.

McQuail, Denis, Mass Communication Theory, 6th edition,Los Angeles, Sage, 2010.

Miltojević, Branislav, Podeljen ekran, Beograd, Filmski centar Srbije, 2011.

Peruško, Zrinjka, ed., Uvod u medije,Zagreb, Jesenski i Turk, 2011.

Saračević, Igor, Pulu otvara potpuno drugačiji hrvatski film, *Tportal.hr*, http://www.tportal.hr/showtime/film/75787/Pulu-otvara-potpuno-drugaciji-hrvatski-film.html (10th July 2013.)

Stojanović, Dušan ur, Teorija filma, Beograd, Nolit, 1978.

Turković, Hrvoje, Narav televizije, ogledi,Zagreb, Meandar, 2008.

Turković, Hrvoje, Teorija filma, Zagreb, Meandar, 1994.

KNOWLEDGE MANAGEMENT

Teaching Universal Decimal Classification (UDC) to Undergraduate Students: A Folksonomy Driven Approach

Tomislav Ivanjko
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
tivanjko@ffzg.hr

Summary

Teaching controlled vocabularies and classifications at an introductory level can be challenging in terms of clarifying theoretical notions and motivating the students. In this light, approaches where students are presented with real life data that illustrates key concepts can be very useful. This case study investigated the difference between subject description of a journal article when using uncontrolled vocabulary (free keywords) and controlled vocabulary (UDC notions) as compared to author keywords, title keywords and UDC notions added by the subject specialist. The study was carried out before and after students attended an undergraduate introductory course to the field of controlled vocabularies and classification systems with a special accent on using the Universal Decimal Classification (UDC). The results confirmed two original hypotheses that UDC driven (post-course) tags will show a greater consistency and have higher agreement with the terms added by the subject specialist than free keyword (pre-course) tags. Since the study was carried out in an educational setting one of the goals of the study was to investigate the possible use of the study to illustrate the basic concept differences between uncontrolled and controlled vocabulary in the light of folksonomies and the use of UDC as a pre-coordinated term list.

Key words: folksonomies, UDC, controlled vocabulary, classification

Introduction

With the rise of digital environments, the use and implementation of controlled vocabulary is the ever more important part of everyday tasks performed by information specialists in the LIS field. Whether searching scientific databases, library catalogues or just Google, basic concepts of the use of controlled vocabularies are always present when the retrieval of information is needed. Knowing the advantages and disadvantages of controlled vocabularies as well as proper

approaches to implementation is at the core of every knowledge organization and representation method.

In recent years, a new wave of user participation in creating and describing online resources instigated a new approach in knowledge representation – folksonomies. Folksonomies rely on the process of collaborative tagging, where many users add metadata in the form of keywords to shared content (Golder and Huberman, 2006), and the totality of these added tags on any different platform forms a folksonomy. From the KO aspect, folksonomies represent a "weak" method of knowledge representation because they don't have means to express semantic relations as traditional methods of knowledge representation such as classifications, thesauri or ontologies have (Peters, 2009) so many authors advocate the use of folksonomies as a complementary method of knowledge organization by using power tags extracted from folksonomies along with controlled vocabularies (Kwan and Chan, 2009; Mendes, Quiñonez-Skinner and Skaggs, 2009). In this notion, folksonomies are useful for providing user warrant, i.e. ensuring that the subject description matches the user vocabulary (Mai, 2011).

Considering the significant impact folksonomies had on the field of knowledge organization, they should be included in the syllabus of any course dealing with controlled vocabularies. One of the first attempts in exploring ways to integrate social tagging into teaching students controlled vocabulary was done by librarians at Boston University's Alumni Medical Library (AML) teaching medical students how to use Medical Subject Headings (MeSH) (Maggio et al., 2009). They created in-class exercises to present MeSH and the concept of a controlled vocabulary in a familiar and relevant context for the course's Generation Y student population and provided students the opportunity to actively participate in creating their education. The librarians found that this exercise – including the pre-class activity, intervention, and post-class evaluation – helped clarify the concept of MeSH, and that couching unfamiliar concepts in the context of popular technologies, can lead to more effective teaching (Maggio et al., 2009).

Implementing a case study

Following this initial approach, a small case study was implemented in the course "Classification Systems", offered to students of the Department of Information and Communication Sciences at the Faculty of Humanities and Social Sciences, University of Zagreb, Croatia. The course is offered at the undergraduate level, holding 6 ECTS credits and is intended as an introductory course to the field of controlled vocabularies and classification systems with a special accent on using the Universal Decimal Classification (UDC) as a practical implementation of a classification system widely used by Croatian libraries. The course is offered each year in the summer semester, and in the academic year 2012/13 a total of 30 students enlisted in the course. There were two main reasons this course was selected for the case study: (1) students enrolling in the

course are at the undergraduate level so they have no prior experience with controlled vocabularies in the form of course credits and (2) their age makes them likely Web 2.0 users so it was expected that the use of social tagging (a part of the Web 2.0 suite) could enable students to better understand the use and purpose of controlled vocabularies and classifications in the LIS field.
This case study had 2 main goals, one strictly professional and the other educational. From the professional point of view we wanted to investigate whether using common pre-coordinated vocabulary (UDC notions) and familiarization with the proper use of controlled vocabulary during the course will have significant impact on the description of the same resource. That meant measuring the term similarity of the pre- and post-course exercise in order to investigate and reflect the possible roles and aspects of folksonomies in the field of controlled vocabularies and UDC in particular. The second goal was aimed at exploring the possible use of these results to illustrate the basic concept differences between uncontrolled and controlled vocabulary in the light of folksonomies and the use of UDC to the students.

Methodology and limitations

Since the course is taught during a single semester, there were a total of 12 lectures and exercises planned that addressed different topics within the course topic during the 15-week period. At the beginning of the semester, during the first exercise, as a part of the lecture on the role of abstracts and keywords, the students were presented with three journal articles from the field of Social Sciences. The articles chosen for description were from a scientific journal, so they had abstract and author keywords assigned as well as UDC numbers added to them from a subject specialist, but were stripped from all of them. Students were given an assignment to write a short description about the topic of the articles and tag them. The exercise took place in an online environment of the Moodle based LMS used by the Faculty where an assignment was created, and the students were not aware that this exercise would be used for a case study. Of the three articles, from the study point of view, first two articles were considered only training exercises, with the results based on the third article planning on being used in the study itself. This was done to ensure that students understood the assignment and were able to use the LMS properly to complete it. After that first exercise, the course was taught during a 12-week period covering the topics of controlled vocabularies, classification systems and the practical use of the UDC. During the last exercise session, a follow-up assignment with the same methodology was undertaken on the fore mentioned three articles. This time students were instructed to summarize and tag the articles again, but this time using the knowledge they acquired during the semester and using the UDC notions as tags. Two hypotheses were:

(1) UDC driven (post-course) tags will show a greater consistency than free keyword (pre-course) tags
(2) UDC driven (post-course) tags will have higher agreement with the terms added by the subject specialist than free keyword (pre-course) tags

Both hypotheses were based on the presumption that using common vocabulary (UDC notions) and familiarization with the proper use of controlled vocabulary during the course will have significant impact on the results and subject approach. After the assignment was completed, it was found that out of 30 students, a total of 25 students completed both assignments so their tagging results of the required article was included in the further analysis. The idea of the study was not to be comprehensive, but to investigate and illustrate specific aspects of the relationship between controlled vocabularies, UDC and docsonomies within the educational setting. Since this study was carried out in a specific setting of an undergraduate course, the small sample of the students is a limitation of the study.

Results and discussion

First, the two tag sets were analyzed using descriptive statistics as presented in the Table 1.

Table 1: Free tags vs. UDC notions

	# OF STUDENTS	# OF TAGS	AVERAGE	# OF DIFFERENT TAGS	TOP 10 TAGS
FREE TAGS	25	146	5.84	66 (45%)	68 (47%)
UDC NOTIONS	25	156	6.24	43 (27%)	101 (65%)

In the pre-course exercise students assigned 146 tags, with an average of 5.8 tags per student, while the in the post-course exercise students assigned 156 UDC driven tags, with a somewhat higher average of 6.2 tags per student.
Further analysis wanted to examine the consistency vocabulary used in both exercises. It was hypothesised that the post-course tags will show a greater consistency because students are using UDC notions and are familiarized with the proper use of controlled vocabulary. This was confirmed by the analysis where it was shown that the docsonomy (all tags of a concrete document, Peters et al., 2011) for the article described with free tags generated 66 different tags (45%) while the one created post-course, using UDC notions had only 43 different tags (28%). The analysis of tag frequency distribution showed that the top ten tags (with the highest frequency) account for 46% (68) of total tags when students were using free tags, while that percentage has risen to 65% (101) when they

were using the UDC notions. When those values are presented on a graph, we can see that both distributions follow the power law (Figure 1).

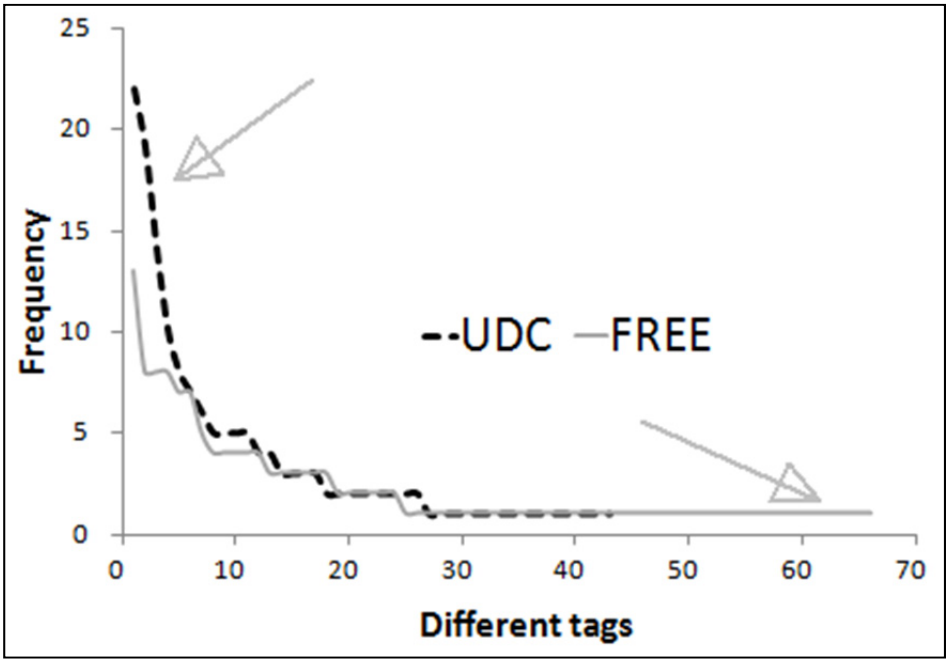


Figure 1: Tag frequency distributions compared

The comparison graph reflects differences where the distribution of free tags produces a longer head and the UDC notion distribution creates a longer tail. Following these results, we can confirm that participants created a more consistent folksonomy when using UDC notions, using fewer tags with higher frequencies. This confirms our hypothesis that the post-course tags will show a greater consistency because students are using pre-coordinated list of UDC notions and are familiarized with the proper use of controlled vocabulary, thus showing UDC driven descriptors more consistent and compact.

Since the article chosen for description was from a scientific journal, it had author keywords assigned as well as UDC number added to it from a subject specialist. From the UDC number added to the article (*004:37.018:007*) UDC notions associated with the number were extracted (9 terms). Added to these two sets, words from the title were also extracted (8 terms) to be used as a comparison set. The idea was to compare added free keywords (pre-course) and UDC notions (post-course) to see the overlap between those sets. As the user set, only tags with a frequency of 3 and higher were used to ensure basic tag validation (added at least by 3 students). Such criteria yielded 18 free tags and 16 UDC notions. The results are presented in Table 2.

Table 2: Term comparison between sets

	NO. OF INTERSECTING ELEMENTS			
	AUTHOR KW (5)	**TITLE KW (8)**	**SUBJECT SPECIALIST (9)**	**FREE TAGS/UDC NOTIONS**
FREE TAGS (FREQ 3+) (18)	2/5 (40%)	4/8 (50%)	1/9 (11%)	3/16 (19%)
UDC NOTIONS (FREQ 3+) (16)	0/5 (0%)	1/8 (12%)	6/9 (67%)	

As we can see from the analysis, when students had to add free terms to describe the article, the highest overlap was between added terms and title keywords (50%). There was also some overlap between free keywords and author keywords, and only one term was the same as that extracted from the UDC number added by the subject specialist. This illustrates several aspects of the free keyword driven crated docsonomy: (1) students presumed that the most important terms are already present in the title of the article so they included them in their description (can be viewed as a copy-paste method); (2) some overlap can be expected between fee keywords and author keywords, since both are uncontrolled and (3) there is a low agreement on terms between student free keywords and those assigned by a subject specialist.

When the same analysis was carried out using the UDC driven set of terms added by the students, the results showed highest agreement with the terms added by the subject specialist, where 6 out of 9 terms were present in both sets. This was again in agreement with our second hypotheses that UDC driven (post-course) tags will have the highest agreement with the terms added by the subject specialist. This was underlined with a low agreement between UDC driven tags and author keywords and title keywords.

The last analysis compared added free keywords (pre-course) and UDC notions (post-course) with a frequency higher than 3. It was shown that there were only three common terms that the students used both in pre- and post-course description. This showed how using pre-coordinated terms can yield very different results from using uncontrolled terms.

Conclusion

This case study investigated the difference between subject description of a journal article when using uncontrolled vocabulary (free keywords) and controlled vocabulary (UDC notions). The study was carried out before and after students attended an undergraduate introductory course to the field of controlled vocabularies and classification systems with a special accent on using the Universal Decimal Classification (UDC). The results confirmed two original hypotheses that UDC driven (post-course) tags will show a greater consistency and have higher agreement with the terms added by the subject specialist than free keyword (pre-course) tags. These results followed the line of some earlier

research (Špiranec and Ivanjko, 2013) showing that, from a knowledge organization aspect, users with more knowledge and expertise could create folksonomies of a higher quality.

Since the study was carried out in an educational setting one of the goals of the study was aimed at using these results to illustrate the basic concept differences between uncontrolled and controlled vocabulary in the light of folksonomies and the use of UDC to the students The methodology and results of this analysis can be used as an illustration of the advantages of using controlled vocabulary in describing resources. It was shown how implementing basic principles of subject analysis and using pre-coordinated terms can yield very different results from using uncontrolled terms. This study illustrated possible weaknesses of folksonomies and the need for complementing them by using controlled vocabularies for enhancing the quality of subject access. This exercise was also useful for illustrating how UDC notions can be used as a pre-controlled vocabulary for subject description. Since teaching controlled vocabularies and classifications at an introductory level can be challenging in terms of clarifying theoretical notions in the field on one hand, and motivating the students on the other, such approach where students are presented with real life data that illustrates key concepts can be very useful.

Further studies in this direction should include larger sample, student evaluation of the approach as well as exploring other ways to implement a case study in the educational process.

References

Golder, Scott A.; Huberman, Bernardo A. Usage patterns of collaborative tagging systems. // *Journal of Information Science*. 32 (2006), 2; 198-208.

Maggio et al. A case study: using social tagging to engage students in learning Medical Subject Headings. // *Journal of the Medical Library Association*. 97 (2009), 2; 77-83

Mai, Jens-Erik. *Folksonomies and the New Order*: *Authority in the Digital Disorder*. // *Knowledge Organization*. 38 (2011), 2; 114-122

Mendes, Luiz H.; Quiñonez-Skinner, Jennie; Skaggs, Danielle. Subjecting the catalog to tagging. // *Library Hi Tech*. 27 (2009), 1; 30-41.

Peters, Isabella et al. Social tagging & folksonomies: Indexing, retrieving… and beyond. // *Proceedings of the American Society for Information Science and Technology*. 48 (2011), 1; 1-4.

Peters, Isabella. Folksonomies: indexing and retrieval in Web 2.0. Berlin : De Gruyter Saur, 2009.

Špiranec, Sonja; Ivanjko, Tomislav. Experts vs. Novices Tagging Behavior: An Exploratory Analysis. // *Procedia - Social and Behavioral Sciences*. 73 (2013); 456–459

Yi, Kwan; Mai Chan, Lois. Linking folksonomy to Library of Congress subject headings: an exploratory study. // *Journal of Documentation*. 65(2009), 6; 872-900.

Facebook: The Good, the Bad or the Ugly?

Mihaela Bingula, Andrea Matić, Vanesa Varga, Nives Mikelić Preradović
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
mihaela.bingula@gmail.com, amatic2@ffzg.hr, vanesa.varga@gmail.com,
nmikelic@ffzg.hr

Summary

This paper presents a student project that used a unique service learning approach to the issue of cyberbullying among elementary school pupils. Students have prepared a combination of lectures (curriculum units) and workshops about cyberbullying that occurs on social networks. Main purpose of the project was to educate pupils about the harmful side of the social networks, violence, and pedophilia. However, students also showed children the variety of ways social networks can be used appropriately in school and how to protect themselves from bullies and predators on the web.

Key words: cyberbullying, elementary school, communication, social networks, Facebook, privacy, protection, anonymity

Introduction

The topic of this project was chosen following a series of bullying incidents on social networks. The authors of this paper believe that children should be gradually introduced to social networks and trained on how to use them safely and in the right way so they can profit from it (education, friendships, etc.). Children should be guided and warned of the dangerous of social networks such as private information, theft and cyberbullying, but they should also be taught how to protect themselves and follow rules enforced by their parents and teachers. It needs to be emphasized that, according to Facebook terms, children younger than 13 should not have a Facebook account. Unfortunately, research already showed that children use Facebook in much earlier age (a survey of kids in U.K. revealed that a quarter of them have a presence on a social network--way below the age limit of 13). The Facebook usage in Croatia is not different from United Kingdom. Therefore, we decided to have children younger than 13 (5th grade) included in our project: not as an encouragement for children to use Facebook, but as a security guide for those of them who already have a "secret" account. Although both parents and teachers can decide to ban social media from their children until they are 13, such approach might be counterproductive to a

child's development, since the ways that young pupils interact and make friends changed a lot during the last decade.

Bullying has been recognized by parents and school officials as a serious problem for decades. But, this harmful behavior has transformed its methods and ways as our communication habits changed. E-mails, text messages and social profile sites are part of our digital era. As reports show, 90% of children are online on a daily basis [10].

Cyberbullying is often described as a new form of traditional bullying. It is "an aggressive, intentional act carried out by a group or individual using electronic forms of contact, repeatedly and over time against a victim who cannot easily defend him or herself" [1]. Or, put differently, it is: "the use of information and communication technologies to support deliberate, repeated and hostile behavior by an individual or group, that is intended to harm others" [3].

Although both bullying and cyberbullying can be defined as harmful behavior, there are some significant differences between these two. Essentially, cyberspace has no physical boundaries. It means that cyberbullying can happen at any place, any time. A victim does no longer have a "safe-place", since cyberbullying operates in the virtual space that can be accessed using the electronic devices.

The other aspect of cyberbullying is anonymity, causing victim to experience even more intense feelings of fear, frustration and powerlessness and encouraging bullies to be more aggressive. Another characteristic of cyberbullying is the potential global audience, unlike the traditional school environment. The distribution of text, photos or videos has never been easier and never harder to erase entirely once it's published online.

To summarize, bullying can manifest itself in at least three different ways [10]: physical bullying, harassment and relational bullying (destroying victim's relationships). Cyberbullying has two main forms [10]: direct, where messages are transmitted from the bully to the victim, and indirect, in which case bully enlists others to bully the victim.

Qing Li [7] quotes Williard's methods of cyberbullying:

a) flaming – sending angry, rude, vulgar messages directed at a person or persons privately or to an online group
b) harassment – repeatedly sending a person offensive messages;
c) cyberstalking – harassment that include threats of harm or is highly intimidating;
d) denigration (put-downs) – sending or posting harmful, untrue, or cruel statements about a person to other people;
e) masquerade – pretending to be someone else and sending or posting material that makes that person look bad or places that person in a potential danger;

f) outing and trickery – sending or posting material about a person that contains sensitive, private, or embarrassing information, including forwarding private messages or images.
g) exclusion – actions that specifically and intentionally exclude a person from an online group.

Cyberbulling can affect academic and social life of victims causing school absence, eating disorders, substance abuse, depression and even suicide.

Methodology

The project presented in this paper *Facebook: The Good, the Bad or the Ugly?* was designed and implemented by three graduate students of Information and Communication Sciences study at the Faculty of Humanities and Social Sciences (who are also authors of this paper) as part of the course *Service Learning in Information Sciences.*

Service learning is a new teaching method rapidly spreading in higher education in United States and Europe [2, 4, 5, 6, 8, 9]. Following this method, students learn and develop their professional and interpersonal skills as well as critical thinking through active participation in structured cooperative activities that address community needs. It also helps universities to contribute to their communities and instills democratic values, increasing the mutually beneficial interaction between universities and the social community.

The project *Facebook: The Good, the Bad or the Ugly?* used a unique service learning approach to the problem of cyberbullying. This service learning project stretched over a semester and took place in the elementary school Julije Klović in Zagreb.

Each graduate student had the opportunity to apply what she learned to "the real world", but also to be a part of the project run by different participants (school teacher, students, faculty mentor and school pupils).

The process of implementing a service learning component involved four separate and distinct stages: (1) Preparation, (2) Action, (3) Reflection, and (4) Presentation, described in the following chapters.

In the project's preparation phase, students identified and analyzed the problem - social network violence and identified a community need - to teach school children about current social network violence issues. Students designed a detailed plan of the project with the university and community members (school teachers and faculty mentor) to ensure collaboration between their University and elementary school Julije Klović in Zagreb.

In the project's action phase students performed all the activities specified in the written project application, while in the reflection phase they produced the results of their reflection: a reflective journal and e-portfolio. The project ended with the group multimedia presentation.

In their 5 years study these students learned many theoretical concepts and applied them to the imaginary or simulated circumstances, but rarely managed to

apply the acquired knowledge to the real world. In this project they had time and will to rethink and implement their own research ideas and they seized the opportunity to transform them into "hands-on" experience, while observing the results and the project's impact to the end beneficiaries.
Finally, during their research on cyberbullying, students came across a great source of information for their project. The Brave phone and Child Protection Center Zagreb expressed their enthusiasm about the project and offered their partnership and support.
The Brave phone ("Hrabri telefon") is an NGO registered in 1997 with the main aim to protect the children from abuse and neglect. Child protection center Zagreb was founded in 2002 aiming to provide effective and systematic support for traumatized children and their families (i.e. abused and neglected children, families affected by the war, etc). From the first contact with the Brave phone team, students experienced their support and engagement in the topic of cyberbullying. They shared students' concern and will to educate children, parents and teachers. With the help of the Brave phone, students also established collaboration with the Child Protection Center Zagreb, which provided them with copies of their handbook on cyberbullying. This handbook was designed for children, parents, teachers and everybody who works with children and wants to protect and educate children and youth about the huge risks of the Internet misuse**.**

Project's preparation phase

In the first stage of the project students worked on identifying a community need, collaborating with the community partner and acquiring new information and skills.
Students have decided to prepare a combination of lectures (curriculum units) and workshops for elementary school pupils about cyberbullying [1] that occurs on social networks. The main purpose of the project was to educate pupils about the harmful side of social networks, violence, and pedophilia. However, they also wanted to show children how social networks can be used properly and how to protect themselves from bullies and predators on the web.
After identifying the need (cyberbullying) and potential community partner (elementary school), the students' focus shifted onto analyzing how to integrate the service learning project into existing curriculum while meeting the educational requirements of the regular school curriculum as well as the pupils' needs.
Students drafted a 10-minutes long presentation at the beginning of the class. But, since the whole class lasts 45 minutes, they decided to devote most of the time to the workshops ideas. Regarding the workshop, students have selected a short animated movie as a motivational instrument for the entire school class.

In the next step, they planned the session of poster design with pupils as a type of artistic workshop, so that these posters can later be used as part of the school exhibition on the cyberbullying topic.
Although students' main objective was to develop curriculum units and workshops that would raise pupils' and teachers' awareness of the issue and effects of cyberbullying in the elementary school, they also wanted to explore opportunities to acquire new skills, to think critically, to adopt new competencies and to test their roles in an environment which encourages risk-taking and rewards competence. The student communicated about the project via Moodle, e-mail and face-to-face meetings.

Project's action phase

In the first phase of the project realization, a student team has chosen the team leader who was responsible for communication with the faculty mentor.
The key goal emphasized in the action phase was to design an interesting presentation and a workshop on social networks and problem of cyberbullying for pupils, but students also planned to make a small exhibition on the topic and include pupils' parents. Considering the different learning preferences of the teammates, the different knowledge level and different understanding of the issue, students had to cope with the existing limitations.
Project *Facebook: The Good, the Bad or the Ugly?* was carried out in two 5th grades and one 7th grade during informatics class at the elementary school Julije Klović in Zagreb.

Introductory and motivational part

As a part of a single lesson (45 minutes), students firstly held a small presentation about violence and potential danger of the online world followed by a workshop about cyberbullying.
The class conducted by students included: (1) an Introduction, (2) the Lecture, and (3) the Workshop. The Introduction consisted of the following elements: (1) an introduction stating learning objectives and what the project was about, (2) background information establishing the importance of the chosen topic and explaining to pupils general terms like communication and Internet, (3) motivational video from Youtube titled: *Znaš li s kim razgovaraš? Online zamka.* ("Do you know who you're speaking with? An online trap").[1] This motivational video is a short animated movie made in collaboration between Center for missing and abused children in Osijek, video club Mursa and elementary school Mladost in Osijek. Drawings and animations were made by eighth grade pupils to warn their peers about the potential dangers of the online world, especially social networks. This powerful multimedia resource helped pupils to get acquainted with the main topic of the project.

[1] http://www.youtube.com/watch?v=HrDXuT2oGFY

The Lecture was delivered as a Power Point slide presentation for pupils and it was divided in three parts. Firstly, students presented the notion of social networks, Facebook and violence in general, addressing issues such as: what to do if you are cyberbullied and why pupils cyberbully. In the second part of the lecture they presented possibilities and ways of online protection and prevention of cyberbullying. Finally, they concluded the presentation indicating positive uses of social networks.

Workshops

The last part of the lesson consisted of workshop activities that actively involved classroom pupils and targeted their multiple learning styles, requiring pupils to implement what they have learned from student lecture and to think about how the lesson applied to life outside of the classroom.

Two sets of workshop activities were developed by students. One set was designed to be used with fifth graders and the other set was designed for use with seventh graders. Specific goals for each group workshop were described in details in the student project plan, as well as all the activities that would take place in the group during the 45-minutes lesson.

Fifth grade students were instructed to write positive and negative aspects of the social networks.

In the first class of the fifth graders, pupils were given red and green cards. They had to write negative experiences and the misuses of social networks on the red cards, while on the green cards pupils had to list the good characteristics of the online communication. On the yellow cards students have written neutral terms like „anonymity" or „availability", relating them to the semantics of online communication, but they also taught pupils about meaning of such terms in both positive and negative context. Red, yellow and green cards were then shaped as a traffic light, showing negative, neutral and positive aspects of social networking.

The other class of the fifth graders had a similar task, writing down positives and negatives, but this time they were given cards of several different colors. They were instructed to write down one good thing about the Facebook on the one card, while on the other they had to write down one bad thing about it. Students made two posters out of these cards: one presenting "good" cards with the „like" button  , and the other containing "bad" cards with the „dislike" button  .

Seventh graders had a bit more challenging task. They were divided in smaller groups of 3-5 pupils and given a task to invent a new Facebook group with a socially useful aim: useful for their class / school or their neighborhood or for some other shared interest. While some groups of seventh graders did an excel-

lent job writing down all the steps of creating such virtual group and describing the possible issues that might occur, other groups were not so engaged.
Results of these workshops were posters which the pupils designed themselves, expressing their reflections and comments about the given topic. Posters were put in the visible place (two in the corridor of school and one in the informatics classroom).
Students have fulfilled the objectives of their project, connecting them with the basic objectives of their study (museology / information science). They have also enabled further communication about the topic of cyberbullying in the whole school, bringing it closer to other pupils who did not have a chance to participate in this project.

Project's reflection phase
During the whole project (preparation and action phase), students were required to write a reflection journal in order to integrate their learning and service experience with the personal awareness and growth.
Reflecting on the service learning process and the skills they mastered, they wrote that these projects represent innovation in education because they provide practical part of it that is usually neglected. The project had a positive effect on their personal development, fostering the sense of personal efficacy and leadership, as well as the ability to work with teammates improving the communication skills and confidence.

Evaluation

In order to receive the feedback from their beneficiaries, students designed a survey that consisted of ten questions. Their goal was to find out the number of pupils with Facebook accounts and if their parents knew about these accounts, since children under the age of 13 are not allowed to open Facebook accounts and can only do so if they or an adult have been untruthful about their date of birth. Students explored whether pupils were familiar with the topic of violence on social networks, whether they experienced violence themselves and if they need and like lectures and workshops such as ones they attended. A total of 49 pupils from two 5^{th} grades and one 7^{th} grade were surveyed. Pupils were asked to answer with "yes" or "no" to questions or rank their claims on a 1–5 scale with 1 being the lowest value and 5 being the highest value. Most of the pupils were male.
Furthermore, pupils were asked if they have a Facebook profile in the first place and if their parents know about it. Students discovered that as many as 68% of pupils have a Facebook account that their parents knew about (Figure1).

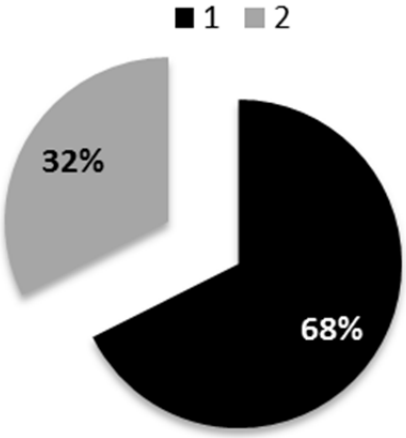


Figure 1. Percent of pupils that have a Facebook profile

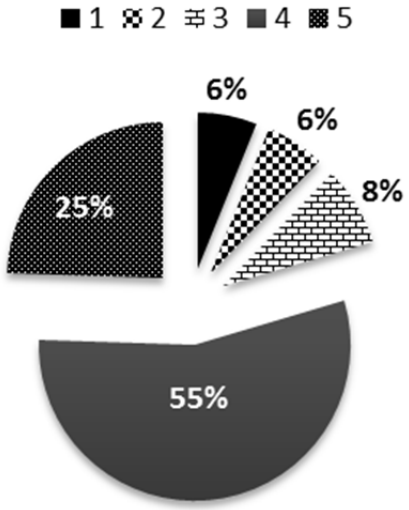


Figure 2. Percent of pupils with previous knowledge about cyberbullying

Also, more than half of the children (55%) were familiar with the topic of violence on the social networks (Figure 2).
Most of the children, 74%, did not experience any problems while using social networks. But, 20% has faced some form of violence, while 6% encountered serious violence on the social networks (Figure 3).

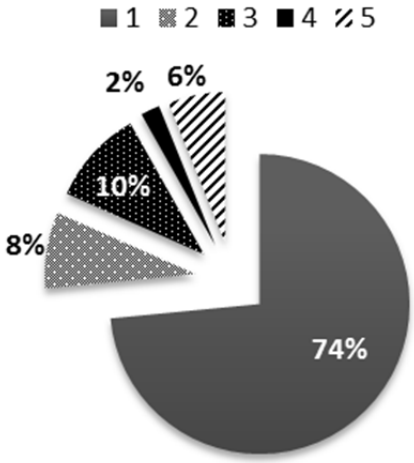


Figure 3. Percent of pupils encountering violence on social networks

Most of the children (53%) admitted that these types of lectures are needed in elementary schools (Figure 4).

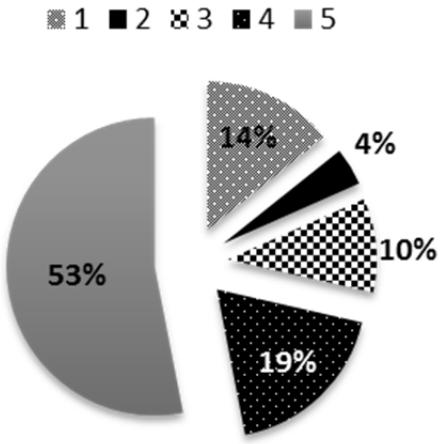


Figure 4. Pupils' attitudes towards lectures on topics such as violence on the social networks

Finally, more than 50% of the pupils declared that they learned something new during the lecture and workshop on cyberbullying (Figure 5). The last survey question allowed them to add their own comments. They acknowledged that there is a lot of violence on social networks, but they rarely dare to talk about it. The survey results have shown that although children encounter violence on the social networks, they tend to be reluctant to talk about it.

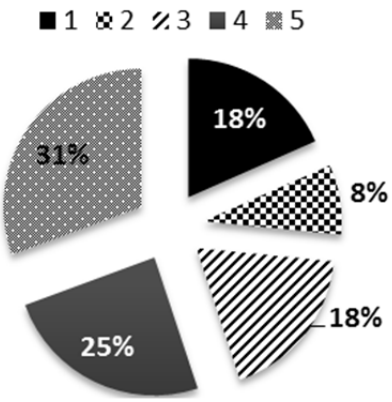


Figure 5. Percent of pupils that stated they learned something new in this lecture and workshop

Conclusion

After the evaluation of their project, students have come to the following conclusions: all pupils (5th and 7th graders) that were part of the project use the Internet and are mostly acquainted with terms that adults use in online communication. Most of the pupils use various social networks, especially Facebook. The majority of pupils are acquainted with the potential dangers of the Internet.

Furthermore, students in this project have revealed an emergent issue in the elementary school curriculum: pupils do not learn about the Internet until they become 7th graders, but they begin to use the Internet much earlier, without necessary skills, knowledge and the quality supervision.

Therefore, students have fulfilled the specified project goals: they have warned 5th and 7th graders about possible danger of using social networks, introducing them to various ways of protection. They have also taught them about the positive use of social networks and encouraged them to participate in the design of posters and small projects.

The service learning project *Facebook: The Good, the Bad or the Ugly?* has addressed a specific community need (violence on social networks) while giving students an opportunity for critical thinking through active participation in a structured activity. For some of the students, this project experience represented a first face-to-face encounter with pupils and the opportunity to teach in front of real pupils instead of their peers. Moreover, students have established a successful collaboration with the community partners: the elementary school Julije Klović, Child Protection Center Zagreb and the Brave phone, as well as with their final beneficiaries - school pupils.

This project brought many long-term contributions to society: it increased public awareness about issues of cyberbullying and about adequate usage of social

networks; it promoted the Internet culture and its benefits to pupils, while the school community witnessed a unique approach to one of the important and contemporary issues in the elementary school.

References

[1] Ahlfors, R. Many Sources, One Theme: Analysis of Cyberbullying Prevention and Intervention Websites, Journal of Social Sciences, Vol. 6 No. 4. 2010, pp. 513-520.
[2] Anderson, D. D. Students and service staff learning and researching together on a college campus. Michigan Journal of Community Service Learning, 9, 2003, pp. 47-58.
[3] Belsey, B. Cyber-bullying: An emerging threat to the 'Always on' generation, 2006. http://www.cyberbullying.ca/pdf/Cyberbullying_Article_by_Bill_Belsey.pdf (Accessed April 19, 2013.)
[4] Cashel, M. L., Goodman, C., & Swanson, J. Mentoring as service learning of undergraduates. Academic Exchange Quarterly, 7(2), 2003, pp. 106-110.
[5] Cleary, C. Steps to incorporate service learning into an undergraduate course. Journal of Experiential Education, 21, 1998, pp. 130-133.
[6] Jones, S. R., & Abes, E. S. Enduring influences on service learning on college students' identity development. Journal of College Student Development, 45, 2004, pp. 149-166.
[7] Li, Q. Cyberbullying in schools: An examination of preservice teachers' perception, Canadian journal of learning and technology, Vol. 34 No. 2, 2008. http://cjlt.csj.ualberta.ca/index.php/cjlt/article/view/494/225
[8] Reed, V. S., Jernstedt, G. C., Hawley, J. K., Reber, E. S., & DuBois, C. A. 2005. Effects of a small-scale very short term service learning experience on college students. Journal of Adolescence, 28, pp. 359-368.
[9] Schaffer, M.A., & Peterson, S. Service learning as a strategy for teaching undergraduate research. Journal of Experiential Education, 21, 1998, pp. 154-161.
[10] Snakenborg, J., Van Acker, R., Gable, R.A. Cyberbullying: Prevention and intervention to protect our children and youth, Preventing School Failure, Vol. 55 No. 2, 2011, pp. 88-95.

Dealing with Information Through Social Lean Thinking

Mirela Grginčić
CROZ d.o.o.
Lastovska 23, Zagreb, Croatia
mgrgincic@croz.net

Krešimir Musa
CROZ d.o.o.
Lastovska 23, Zagreb, Croatia
kmusa@croz.net

Summary

Information can be managed in different ways. This paper explains how to increase effectiveness through information governance and to introduce continuous business process improvement into any organization with help of lean thinking through social business environment.

Key words: lean, continuous improvement, social business network, knowledge

Introduction

This paper is not focused on explaining what lean means or what the best lean practices are, but instead on how to introduce lean management in a unique way into any organization that wants to improve information and knowledge sharing between their employees. One of the important aspects in information management is the idea management and this paper will explain why social collaboration platforms are becoming inevitable today regardless whether an organization is focused on production or on services industry.

New approach to software support for lean transformation

Although there are experts who think that lean management does not have anything to do with information technology, it can be proven that each principle (lean, continuous improvement, kaizen etc.) can be very well supported with IT if the IT system is completely in accordance with defined rules and practices. Not only this, but also well-designed and flexible IT system can bring a higher value for the customer than just following lean management best practices without IT support. Today, we can see lean initiatives not only in manufacturing industry, but also very often in service organizations and we can say without doubt that lean thinking can be applied to all business activities, especially information governance.

Through experience with customers, who either plan or already have implemented lean, we discovered a lack of IT support in their continuous improvement process. One possible solution to this challenge can be a simple idea management tool. There are many existing idea management tools that are not in accordance with lean principles nor applicable for every organizational process. Therefore, CROZ has developed a new approach in lean management; IT support which combines three key elements for a complete and rounded solution no matter on which platform it is delivered:

- Social business collaboration platform
- Lean management and continuous improvement principles
- Business process management

Social business collaboration platform features can help understand why choosing IT support should be the next step after recognizing the need for continuous improvement in an organization.

Social networks started to grow globally after 2005 when Facebook opened to anyone over age 13 and Twitter was launched. In the years that followed, people started using social networks as everyday communication platforms in private life and business as well. In 2011, LinkedIn had more than 90 million professional users. At the same time, IT companies that were dealing with enterprise systems recognized this trend and started releasing their enterprise social networks intended for internal employee's collaboration in organisations (IBM Connections, Jive, Yammer etc.). In the last few years, companies are starting to realize many benefits of such platforms and tools aimed for improvement of their internal communication and information management.

Idea management solution on top of social platform can be considered as a central place where suggestions on continuous improvement are gathered and further developed in business process management tool.

How does social business collaboration platform fit into lean?

Information governance means dealing with an effective and efficient use of information in organizations. Social business collaboration platform has the same purpose; to find and provide information in more effective and efficient way than through e-mail, usage of static document management systems, etc. First among the lean principles collaboration platform should support is "coaching people".

Lean considers that each lean manager who is an expert for his/her area needs to dedicate his/her time to teach and coach other people. Social business collaboration platform can be very useful for knowledge distribution and easier information access. Each time when a new employee is hired, manager does not have to waste time on starting from the scratch with the new employee. All the knowledge is kept on collaboration platform and does not disappear.

Prerequisite for making excellent products and provision of quality services is coaching people. According to lean it is better if manager acts as coach rather than director. This approach shows people that they are able to approach their teachers without fear of failure or denial.

Associate knowledge with people, not documents

Most organization keeps their knowledge in document management systems. The problem with knowledge in static documents is that it gets old very fast (knowledge changes dynamically). Thus, it does not represent a natural way of sharing because people in real world share information in real time, mostly through different updates, posts, blogs, wikis etc.

If expertise in certain area is needed usually people do not search all possible documents on specific topic. Instead, an expert should be found first and his/her documents and blogs. Social business networks are focused primarily on people and their expertise because knowledge needs to be always fresh and relevant.

Continuous improvement through idea management

In further text, focus is given to Idea management and continuous improvement of business processes throughout information lifecycle: from idea creation, its transformation, up to idea implementation. Idea management that is implemented in accordance with lean principles can also support better management of information in organization.

Idea management in lean organisations

In order to give their best ideas, employees need to be challenged and have opportunity to express what they want without fear of bad feedback. Respect for people does not have to be achieved only through awards and good manager-employee relationship. There are many factors that can influence the level of seriousness and enthusiasm with which employees will get to work and understand their role in an organization. As mentioned, one of the factors is to feel safe and brave enough to express their opinion and have motivation to improve their work.

Lantech is a company researched by Womack and Jones [4] and can serve as a good example of challenges that company needs to overcome when implementing lean thinking. One of the main challenges for the Company at the first stages was how to consolidate communication barriers between different departments, especially the information flow. In this sense, collaboration and feedback from other employees/managers needs to be considered very important in the process of breaking down the barriers.

In order to solve existing problems, they need to become visible first to all shareholders included in a specific business process whether these are employees, managers, board of directors, etc. Furthermore, visibility needs to be enabled in order to fill social network with relevant content. Through visual man-

agement of information this can be easily done and checked because many people are included in the process.

Sharing information in a social way

Not giving people an alternative for sending e-mails basically means forcing them to do a wasteful work. Social business networks can enhance teamwork through applications such as Wiki, Blog, and Communities etc. (see Picture 1). This is especially useful for geographically dislocated teams. Social business platform should at least have the following functionalities in order to achieve quality information governance:

- Possibility for teams to work on a project in different, customized and separated environments
- Forums for discussions about business topics in non-formal way in order to motivate people to give their opinions and answers to others questions
- Blogs for publishing new information, e.g. manager articles with company news and strategy, employees blogs on specific business domain, etc.
- Files sharing and versioning for easier file distribution
- Possibility for a whole team to work on one document (Wiki)
- Idea management functionality to enable continuous improvement through employees idea sharing
- Profiles with "wall" feature
- Expert network that can be searched by many criteria in order to find information and experts easily

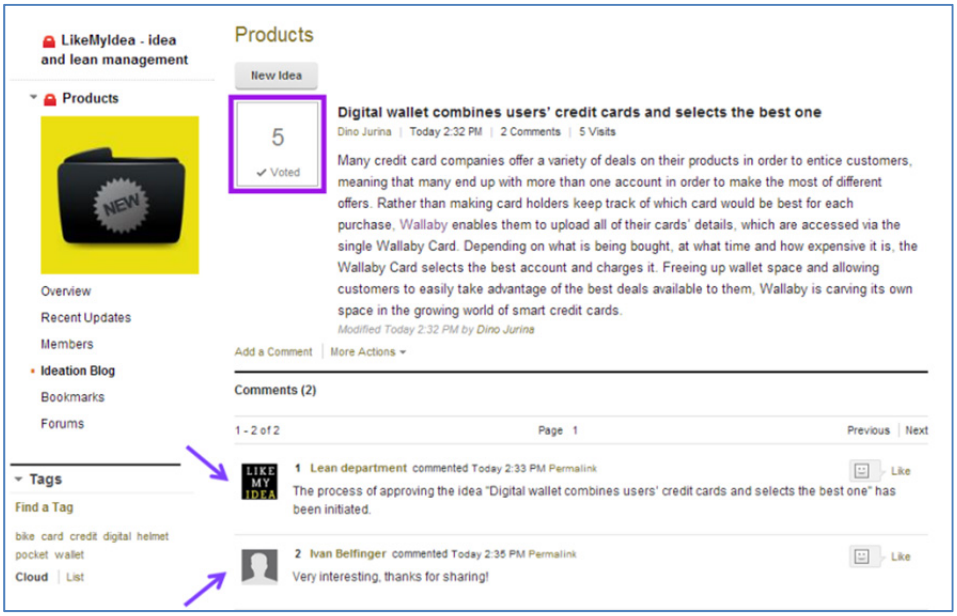
Gathering ideas from employees for continuous improvement process

Management usually gives a task to employees to share their ideas on a specific topic that is in their interest (not important if it is via e-mail, web form, in person or written on a paper). Depending on whether the task is obligatory or -optional, two scenarios will probably happen. If the task is obligatory people will do it only because they have to and it is very unlikely that quality and innovative ideas will be gathered. If the task is optional, most employees will not even try to participate, because sending an e-mail or even worse, saying the idea in person to their manager frightens them.

The approach should be different. Managers should not ask for ideas only on certain topics or when they think there is a problem that needs to be solved. People should have a chance to express their opinions about the things they personally consider important at the time. These ideas often will not have anything to do the with the business domain.

The way ideas are given is also important. If the ideas are sent only to manager's e-mails, managers will have a lot of individual ideas. Between all these ideas (especially if the organization is big) there will be individuals with similar

ideas or perhaps with small differences that can decide if the idea is feasible or profitable for an organization or not. In the first scenario, these individuals do not know that they have given similar ideas and cannot work together to make a single excellent idea. The key is to enable people to work together and to see all proposed ideas in order to comment each other's ideas and improve them.



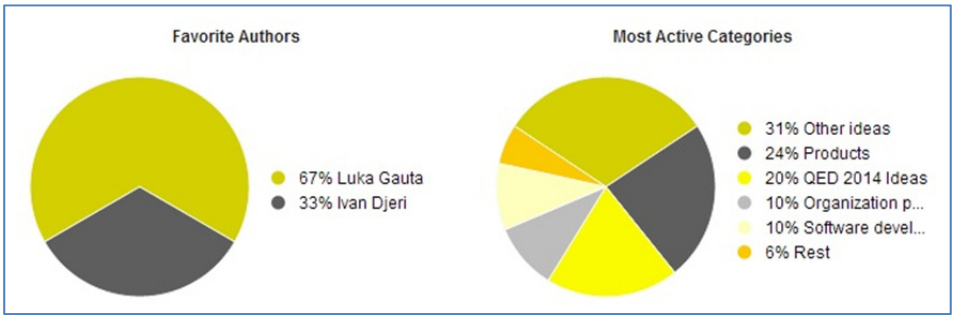
Picture 1. Commenting and feedback on the ideas provided
(Source: LIKE MY IDEA tool)

It is a waste of time if lean department leaders need to look through e-mails in order to read every single idea. In big organization, this is not even viable. Why not allow the whole organization to evaluate the ideas through a simple voting process instead? Lean department can decide on the number of votes necessary for an idea to reach the next step of approval and deal only with ideas that are already filtered and improved among employees.

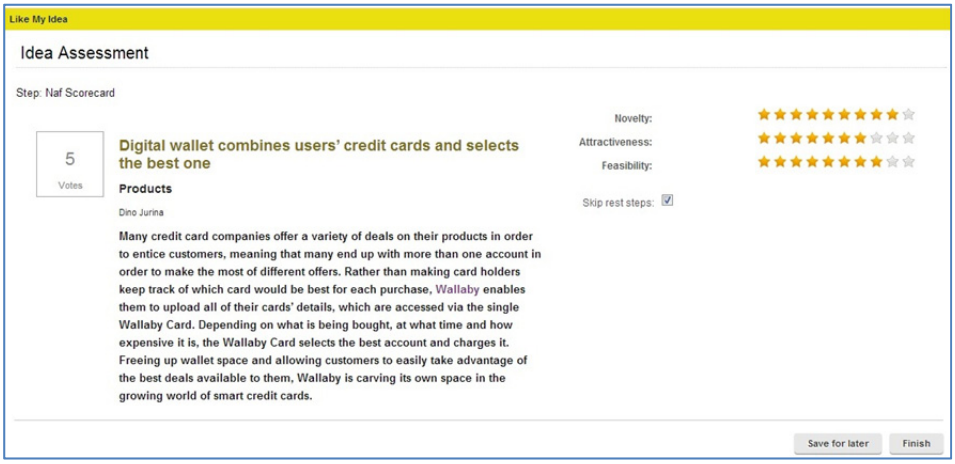
In the Toyota idea management system continuous improvement and innovation initiatives guarantee that team member's ideas will be heard and acted upon. LIKE MY IDEA also implements the system of trust among employees.

Continuous improvement process is possible because of following lean waste elimination: elimination of paperwork, no overloading of coordinator because only best ideas get to the evaluation process, electronic notifications and task assignments, escalation, visual management and metrics.

If we simplify the process we can say that continuous improvement is conducted through following449juYx0ni steps: empower the employees, measure and evaluate ideas and the last one is to improve or act upon the ideas.



Picture 2. LIKE MY IDEA metrics
(Source: LIKE MY IDEA tool)



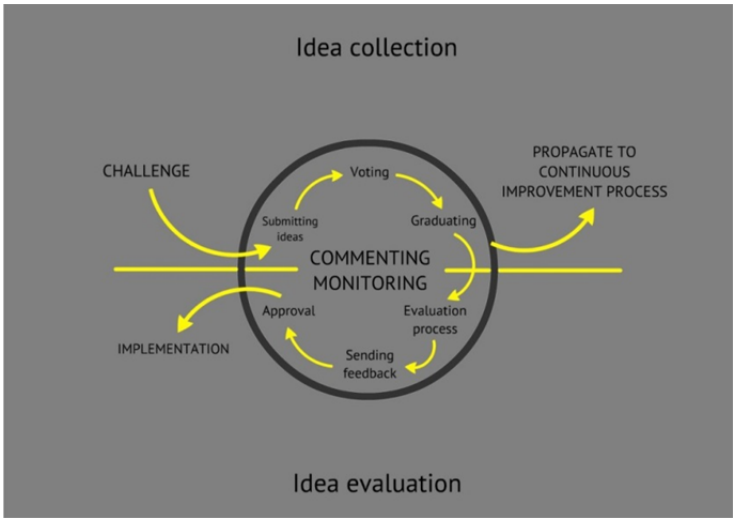
Picture 3. First step in idea approval evaluation
(Source: LIKE MY IDEA tool)

There are many advantages of implementing such approach of communication on social business platform:

- Employees are more willing to give ideas because of the less formal environment, e.g. the ideas can be given in form of Blogs
- Ideas are visible to everyone and others can give constructive comments on already proposed ideas, and these comments are not considered as criticism, but rather as a help to improve the idea
- Each vote has very positive influence on the idea creator
- Gamification can be used on social business platform in order to motivate employees to participate (employees collect badges for voting, proposing ideas or commenting)

Achieving flexibility with business process management platform

Employees have the means to create ideas on social business network in a motivating and informal environment. Ideas can be given both to general continuous improvement topics and also as a response to certain management challenges. If there is no need for all employees to participate in some topics, they can be closed for smaller groups or specific departments.



Picture 4. Idea collection and evaluation process
(Source: LIKE MY IDEA, http://www.likemyidea.com/like-my-idea/lmi-overview/)

After idea is given, the voting process takes place. All employees who have the option to give idea also need to have the option to vote for the idea or comment it. When enough votes are collected, the idea is ready to be propagated for evaluation. Different criteria, scorecards and other methodologies can be used for idea evaluation. The way the idea will be evaluated depends on customer practices. Therefore, IT solution that supports lean thinking in any organization should be very flexible and adaptable to customer's specific business process. There is no single implementation that can fit each customer's need. That is the reason why business process management platform is inevitable in IT supported lean organization.

Business process management platform enables very quick and flexible adoption of approval/evaluation/implementation process. The process can have three simple steps for organizations with smaller lean departments or very complex procedure that is impossible to change dynamically without business process management tools.

Although the procedure can vary from customer to customer, one step should never be avoided. All parts or at least most of the parts of the approval process should have output in form of feedback from lean department that is sent back to social business network (e.g. in form of comments on idea). This step is very important because of psychological reasons. This way, employees can see that something is happening with their ideas (e.g. they could see that their idea is now in redesign phase or that it needs an additional input from the author). Although this seems as a small step in the process, it has a big influence on employees’ motivation.

Conclusion

Through application of features and principles mentioned in this paper, idea management tool can grow into creative and intelligent community where employees will try to continuously improve their work and help others in doing so.

References

CROZ d.o.o., LIKE MY IDEA, 03.07.2013, http://www.likemyidea.com/like-my-idea/lmi-overview/

Dukovska-Popovska, V. Hove-Madsen, K.B.Nielsen, Teaching lean thinking through game: some challenges, 25.09.2013, http://www.sefi.be/wp-content/abstracts/1092.pdf

Kovacheva A.V., Challenges in Lean implementation – Successful transformaton towards Lean enterprise, 25.09.2013, http://pure.au.dk/portal-asb-student/files/9093/ak83188...pdf

Larman, Craig, Bas, Vodde, Lean Primer, 03.07.2013, http://www.leanprimer.com/downloads/lean_primer.pdf

Ries E., The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses, 2011.

Rother M., Toyota Kata - Managing People for Improvement, Adaptiveness and Superior Results, 2010.

Womack J.P and Jones, Beyond Toyota: How to Root Out Waste and Pursue Perfection, 25.09.2013. http://www.pcb.org.za/upload/files/beyond-toyota.pdf

APPLICATIONS FOR E-SOCIETY AND E-GOVERNMENT

Creative Business Environment as a Factor in Improving the Effectiveness of Maritime Agents

Mila Nadrljanski
Faculty of Maritime Studies – Split
Zrinsko Frankopanska 37, Split, Croatia
milamika60@yahoo.com, mila@pfst.hr

Veronika Domitrović
Visoka škola za inspekcijski i kadrovski menadžment u pomorstvu
114. Brigade HV-a 12, Split, Croatia
veronika.domitrovic@icloud.com

Sanja Frkić
OŠ kraljice Jelene
Put mira bb, Solin, Croatia
sanjafrk@gmail.com

Summary

This paper presents analysis of a survey conducted on marine agents and agencies in Split. The purpose of this research was to evaluate organizational effectiveness based on creativity of maritime agents and to determine if a creative attitude increases the level of organizational creativity in practice of maritime organizations in general. The research was empirical and survey was based on SOQ and factor analysis. The work of maritime agents is an important economical factor in a maritime-oriented country such as Croatia.

Key words: organizational creativity, maritime inspectors, maritime agency

Introduction

The purpose of this research was to evaluate organizational effectiveness based on creativity of maritime inspectors. A successful company is the one whose business is constantly growing and improving. A high challenge climate exists when the people are experiencing joy and meaningfulness in their job, and therefore, they invest much energy (Isaksen & Lauer, 1998).

Creativity is a complex and multi-faceted phenomenon. A general framework that approaches the more comprehensive understanding of creativity has been the classic four P's (Person, Process, Product and Place). For example, the most comprehensive picture of the creative person can be drawn by considering not only the characteristics or traits of the person, but also the kind of environment

or context in which the person works, the kinds of mental operations used, as well as the nature of the desired outcomes or products (Isaksen & Aerts, 2011). Organizational effectiveness and functionality greatly depend on the creativity of the workforce.

The survey for marine inspectors

This survey was based on Situational Outlook Questionnaire (Isaksen et al., 1999).

1. Do you consider business as a challenge and personal involvement? Are you encouraged, emotionally involved and committed to work?
2. How free are maritime inspectors to decide on the method of performing their tasks? Do they have time to consider things before acting?
3. Is there dynamism in your work (plenty of events in organizational life)?
4. Is there trust and openness in your work? Do people feel free to express their opinions and offer different viewpoints?
5. Is there enough time to elaborate new ideas? Do you have time to explore and develop ideas which were not a part of the original task?
6. Is there playfulness and humour? Is the atmosphere relaxed? Is easy-going communication considered acceptable?
7. Are there any conflicts? To which extent do people engage themselves in interpersonal conflicts?
8. Is there a support to new ideas and are there means for testing them? Do people listen to each other generously?
9. Are there any debates? To which extent do people engage themselves in lively discussions on real interest issues?
10. How much risk is acceptable? Is failure considered normal?

Empirical research

Based on factor analysis, the dimension "dynamism" was not included, and a questionnaire was reduced to nine dimensions. This instrument contains 50 items, and is constructed to evaluate supporting of creativity and change of every particular context. The subject of this research is the measurement of business creativity and development of creative environment in the work of maritime inspectors. The problem of this research can be formulated by questions: How effective is marine inspector who fosters creative business environment? Can fostering of creative organizational environment in maritime organizations result in increasing of organizational creativity in general?

Answers to these questions are aimed at developing knowledge about variables that allow the operationalization of creative business environment as a factor in improving maritime agents' effectiveness.

The objectives are achieved through the following tasks:

- determining inspectors' attitudes and their latent structure towards creative organizational environment;

- determining the correlation between marine agents' social status (sex, experience, position, level of education ...) and their attitudes towards the creative organizational environment;
- discussion about the results and drawing conclusions relevant to improving the effectiveness of management by fostering creative organizational environment

Research hypotheses

Ha_1: The level of business and organizational creativity of maritime agents contributes to the improvement of the overall creativity in maritime management organizations, as well as to the improvement of creativity in maritime business.

The type of the research

The research is empirically based. A basic methodological approach used in the research was survey method[1] in which data were collected and analyzed with appropriate statistical method, presented and discussed. This method was chosen, as it proved to be quite suitable on the field for fact-finding with this type of variables. The method of theoretical analysis was used as a supplementary one. The contents of scientific debate and research dealing with these issues were also analyzed.

Population and sample

The sample was meant to be drawn from the population of all maritime agents in Split. The sample was identified by random drawing using the systematic sampling method (every Nth from the list), providing that the original sample totalled in xx companies (it was expected that 20 - 30% of companies would not provide valid survey material). There has been a final sample of yy companies with valid survey material. The survey was being conducted from September, 2009 to January 2010. The confidence level of $p < 0.05$ was estimated in the formed pattern.

Variables

The main data collection technique was the electronic survey and personal survey (by authors of the paper). Sections in the questionnaire were:

- research unit and subject identification,
- questions identifying creative environment in the organization,
- questions on subjective preferences, satisfaction and attitude towards life

[1] Answering questions about feelings, behaviour, etc. especially given the time variable (the change of a specific populations over time).

Data Processing

Data were entered into the basic matrix and subjected to multiple controls (logic control, normality, regression analysis demands, item analysis). All variables were subjected to bivariate analysis. According to the prevailing nature of the data the selection of statistical methods was made that corresponds categorical data:

- Non-parametric tests - Chi Square test,
- Logistic regression,
- Path-analysis.

Depending on the nature of data, some other methods were applied (graphical representation of variables for exploratory insight etc.).

Survey results

Variable system reliability

Before the data analysis, it is necessary to determine to what extent the unit of variables "affects" the underlying construct of the research, namely, the construct of organizational effectiveness and the construct of creativity. Due to these checks, an item analysis was performed, using the reliability test by calculating Cronbach's alpha.

The high value of Cronbach's alpha (Cronbach's alpha: .966550) indicates the high reliability of the variables used for the underlying constructs of this research.

The average correlation between variables is low ($r = 0.00$), allowing discrimination to identify relations between individual variables. Detailed analysis of individual variables, their contribution to the reliability of the total variable sample, suggests the existence of variables which increase the level of Alpha (a total of 33 variables) when excluded from the variable sample. There are also variables whose exclusion from the total variable sample causes low Alpha level (77 variables). The first category of variables has a lower value for this cognitive content, while the second category has a high value.

Characteristics of subjects

Before analysing the survey results, it is necessary to determine the reliability of conclusions based on statistics that represent normal distribution of variables. If it proves that the appropriate distribution of data from at least ordinal measuring scale is normal (Gaussian curve[2]), then we can apply more powerful parametric tests. If this is not the case, it is necessary to apply less powerful parametric tests. The results of parametric tests applied to the data that do not satisfy the

[2] Gaussian distribution or "normal distribution" is an important family of continuous probability distributions, applicable in many fields. Each member of the family may be defined by two parameters, mathematical expectation and variance (dispersion) σ^2.

premise of normality, can be taken as a guide to further analysis, but not as a reliable basis for inference.

Demographic analysis of subjects
The requirement of normality is satisfied by the variable *Work experience (RDIS8)*. Other variables do not meet this requirement and this must be taken into account when choosing statistical tests. The absence of normal distribution for most variables limits the reliability of conclusions about subjects' uniformity of attitudes, as confirmed by coefficients of variation (except for variable STRSP6, where deviation of variability from average values is over 50%). The ratio of 40:60 in favour of the male subjects may be considered satisfactory for the social circumstances in Croatia. Data indicate that 90% are of urban origin and 70 % have high school education. Working experience shows that most subjects belong to the younger (51%) and middle (48%) generation.

Attitudes towards professional development factors

Dimensions of Professional Development

When asked what dimension is considered more important for professional development, there were three answers offered:

1. subjective - ability, self-confidence, interests
2. objective - working conditions, promotion, jobs and positions that vary in complexity, responsibility and organizational level
3. they are equally important

Respondents generally attributed equal importance to subjective and objective dimension (three quarters of the responses). Such a response can be interpreted as the lack of understanding of dimensions and their importance. Concerning respondents who gave priority to one of the two dimensions, the relatively higher proportion prefers subjective assumptions of professional development (15%) and a smaller proportion prefers priority to objective factors (9%).

Characteristics of subjects relevant to successful professional development
Subjects were asked about the most important qualities for a successful professional development. They had to circle three characteristics among:
high IQ, nice physical appearance, male / female, social origin and status of their parents, friendliness, obedience and resentment, moral integrity and independence, work, expertise, favouritism and initiative.
The results revealed three priority characteristics for successful professional development:

1. High IQ – 37.4%
2. Work – 44.9%
3. Expertise – 49.2%.

They were followed by moral integrity and independence (22.1%) and initiative (26.8%). The least important were features: male / female, social origin and

status of parents, and favouritism. Based on this ascertainment it can be concluded that the majority of respondents prioritize characteristics subjective to change if the individual is motivated and unless cultural constraints prevent the implementation of the necessary changes in individual's behaviour.

Willingness to engage in improving of professional development
On the question about willingness of intense commitment (to work, further training, travelling ...) in order to enhance their professional development, three quarters of respondents replied that they are quite ready and completely ready. These results lead to a conclusion that the activities of maritime inspectors are open to learning, which is essential in developing creative organizational environment.

Correlation analysis
Correlation analysis indicates that the higher the education, the lower the willingness in making effort to improve professional development (rsp = 0.147 with $p<0.012$). Furthermore, respondents who spent their childhood and youth in more urban environment, showed more willingness in additional efforts for professional improvement (rsp = 0.122 with $p<0.037$). The same relation was observed to be correlated positively on the levels of school education and the willingness to make additional efforts in professional improvement (rsp = 0.123 with $p<0.037$).

Respondents' satisfaction with the professional achievement and conditions
Four questions were asked regarding their satisfaction with the achievement and conditions of professional development: a) about their *own* professional development, b) about general social conditions for development, c) about attention paid to developing and motivating staff, d) about the possibility of acquiring new professional skills through different educational programs.
Concerning satisfaction with their own professional development, 50% of subjects showed the highest positive attitude (the average level of 3.93 with relatively smallest CV[3] = 22.13%), the other 25% showed satisfactory attitude and only 6.2% showed dissatisfaction. The opposite relation was indicated towards general social conditions for professional development of maritime experts (average level of 2.86 with relatively greatest CV = 39.51). In this case the proportion of respondents who indicated dissatisfaction with social conditions was much higher (36.1%) than those who expressed satisfaction (28.7%). Respondents' satisfaction with the situation in organization, especially with employers' attitude towards developing and motivating staff, is in the central position on the satisfaction scale (the average level was 3.42 with a CV = 33.04%). It

[3] Coefficient of variation.

should be noted that a significant percentage of subjects was undecided or dissatisfied (47.7%) with the evaluation of personnel management in their organizations. The same can be stated for the level of respondents' satisfaction with the possibility of acquiring new professional skills through different educational programmes (average level of 3.40 with CV = 33.52). However, the number of subjects who were satisfied with the possibility of further education was 51.4%.

Correlation analysis

High positive correlation was observed between the variables "Satisfaction with attention paid to developing and motivating staff" and "Satisfaction with the possibility of acquiring new professional skills through different educational programs". A strong correlation between variables "Satisfaction with their own professional development" and "Satisfaction with attention paid to developing and motivating staff" was also observed. The respondents perceive the attention to development and motivation of maritime agents as an important factor in creating environment of satisfaction.

Personality characteristics of respondents

Respondents were asked ten questions with implicit introspective assessment of the personality traits:

The highest ratings and level of mutual consent (CV) respondents indicated responding on two variables: "I find the tasks or activities that require autonomy and responsibility suitable for me" (81.6%); and "Demands that I impose on myself are within the limits of what I can do. (82%)

Orientation on activities relying on the abilities and making efforts identified 73.5% of respondents.

About three quarters of respondents (75%) identified introvert attitude, tendency of self-valuation (without comparing their effects and success with others). Tendency towards increased effort in solving problems at work identified 77.6% of respondents. Self-control in a situation when the job requires performing of annoying, repulsive and boring tasks identified 68.9% of the respondents, and more than one quarter did not identify with this feature (28%). Another situation of having self-control with the task of no interest for respondent was identified by a high percentage of respondents (71%). But 27.8% of respondents did not identify with that kind of self-control.

Issues implying negative attitude towards the persistence of problem solving, contained in the variable "Dealing with one problem for some time is a waste of time" divided respondents into two approximately equal subsets: 49.9% of those who identify with the reluctance of long-term problem solving and 48.7% of those who identify positively with it. The same variable was set in the positive sense "I am attracted to demanding and long-term tasks", and resulted with 57.6% positive, and 45% negative identifications. Finally, the variable "I am indifferent when I do not fulfil my obligations" was identified negatively with

75.9%, indicating respondents' refusal to identify with indifference to the commitments. This issue also raised the largest discrepancy with relatively highest coefficient of variation (63.40%).

Correlation of subjective characteristics

Correlation analysis within subjective characteristics of the respondents shows the strongest positive correlation between variables "I do my best even if I am not interested in a subject" and "I work on problem solving, even though it requires a greater effort". Variable "I'm attracted to tasks or activities whose outcome depends on my ability and effort" has a strong correlation with other variables of subjective characteristics of the respondents:

- Concerning my success, I prefer satisfying my own internal criteria than comparing to others.
- I work on problem solving, even though it requires a greater effort.
- Demands that I impose on myself are within the limits of what I can do.

Creativity analysis demanded grouping of subjects into categories according to values attributed to all views / claims calculated from the scale applied in the questionnaire. The variable CRELEV was obtained by calculating the sum of all values attributed to each subject related to creativity. However, this variable measures respondents' attitudes to creativity within the organization where they work. This study analyses level of creativity of organization, so it was necessary to find an indicator of creativity based on the average score of all survey respondents within the same organization.

For that purpose, we calculated average values of more respondents in the same organization (for organizations with one respondent, the individual values were retained). Variables CREORG and CREORG_rnk were obtained this way, the latter meant organization ranking as compared to the average value of the variable CREORG (where 100 is the average value of all organizations).

Concerning the average level of creativity grade of all respondents was 0.396 (median 0.428), we can conclude that in the sample of organizations in the research, the organizational creativity is at the "zero" level, and that ideas and behaviours of creative organizational environment are at early stages.

The analysis has led to the realization that there is a correlation between organizations and the level of organizational creativity. This correlation is strong with the weak negative correlation. Lambda test suggests that organization makes more influence on organizational creativity than vice versa. This finding is significant because it indicates an importance of a maritime inspector in creating a creative organizational climate. Such an atmosphere is not created by the spontaneous initiative of employees, but by initiative and interest of employers. Creative relationships are the key to success and, according to Nadrljanski Đ. and M. (2013, p.54), success is a resulting feature and has to be related to management of interactions rather than management of actions.

We verified these expectations by testing connections between grouped variables of organizational creativity and subjective characteristics of respondents. Tests of connectivity of all the variables from the battery of subjective characteristics of the respondents confirmed that only school education (SCED4[4]) was significantly associated with block variable CREORG4r at the level of organizational creativity.

The test of the power connection points to its low level. Pearson's correlation coefficient[5] shows a weak negative correlation between the observed variables. Lambda test of association between these variables confirms characterizing it as the dependent variable (the error of reasoning is reduced to 9.4% versus 0% if the variable SCED4 is put in that position). This finding suggests that school education is an important stimulus for creative behaviour.

Relatively low level of respondents' school education indicates the tendency in accepting creativity, but there is no required interest in management structures of activating such a tendency. More detailed insight into the connection between the group variable CREORG4r[6] with dimensions of creativity, confirms above mentioned lack of initiative in creating and developing creative organizational environment.

Selecting variables contributing to the level of organizational creativity we can distinguish the variable formulated with a statement "Conspiracies, traps, struggle for dominance and territory are common elements of life in the organization," which describes the environment opposite to one postulated as a creative organizational environment. Only two variables belong to a subset of positive claims while others are negative.

Conclusion

Results of this study show that marine inspectors are mainly unsatisfied with current organizational practices and are therefore often oriented to modern approaches of improving effectiveness. Their tendency in accepting creative environment is positively correlated to the level of effectiveness of management in general. Social and experiential qualities of maritime agents are part of their attitudes and determinations towards fostering a creative environment. Finally, personality traits proved to have a significant impact on their attitudes towards creativity. Concluding the scope of this research one has to consider its limitations. There were certain limitations when determining the sample, as a number of agents and agencies did not respond to the survey material. To compensate

[4] Educational level of Maritime Agents

[5] Pearson's correlation coefficient (r) is the linear relationship measuring between two variables in a sample. The value of r^2 is typically taken as the percent of variation in one variable explained by the other variable.

[6] CREORG4r is short for creativity of the organization where maritime agent works.

for this limitation, we used the request triangulation in analysing the results. The results of this research confirm that the development of maritime organizations is inclined to learning, and their advantage over conventional Tayloristic organizations is creativity, openness and change.

References

Rosenberg, Marc Jeffrey. E-Learning: Strategies for Delivering Knowledge in the Digital Age. New York: McGraw-Hill Professional, 2001

Horton, William; Horton, Katherine, E-learning Tools and Technologies: A consumer's guide for trainers, teachers, educators, and instructional designers. New York: John Wiley & Sons, 2003.

Bullen, Mark; Janes, Diane P.. Making the Transition to E-learning: Strategies and Issues. Hershey, PA.: Idea Group Inc (IGI), 2007.

Bates, Toni (2001). National strategies for e-learning in post-secondary education and training. Paris: Fundamentals of Educational Planning – No. 70, UNESCO, 2001; Http://unesdoc.unesco.org/images/0012/001262/126230e.pdf (Access 2012)

Isaksen, Scott G.; Aerts, Wouter S.. Linking Problem-Solving Style and Creative Organizational Climate: An Exploratory Interactionist Study. // *The International Journal of Creativity & Problem Solving.* 21(2) (2011), pp. 7-38

Isaksen, Scott G.; Lauer, Kenneth J. The Relationship Between Cognitive Style and Individual Psychological Climate: Reflections on a Previous Study. Buffalo: Creative Problem Solving Group, 1998; p. 5

Isaksen, Scott G.; Lauer, Kenneth J.; Ekvall, Göran. Situational Outlook Questionnaire: A measure of the climate for creativity and change. *Psychological Reports, 85* (1999) pp. 665-674

Nadrljanski, Đorđe; Nadrljanski, Mila. Teorija sustava i upravljanja. Split: Redak, 2013; p. 54

Paulsen, Morten Flate. *Megatrends in e-learning provision: Preliminary project report on Sweden Finland, Denmark, Germany, and the Netherlands.* Brussels: Presentation at EuNeOn founding ceremony (2006)

Ministry of Advanced Education. (2006). Degree Program Review Criteria and Guidelines. British Columbia, Canada. http://www.aved.gov.bc.ca/degree-authorization/documents/degree-program-criteria.pdf (Access 2007)

Proposing the Model for Croatian Remote Access Safe Centre for Statistical Microdata

Martina Poljičak
Croatian Bureau of Statistics
Ilica 3, Zagreb, Croatia
poljicakm@gmail.com

Hrvoje Stančić
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
hstancic@ffzg.hr

Summary

The data collected in National Statistical Offices (NSOs) present a valuable source for the research community. These data, when put to some secondary analyses, bring new benefits for society. The growing need for researchers' easy access to publicly funded data is recognised in the European Statistical System (ESS). There are on-going efforts for establishment of pan-European network for Remote Access (RA) to Microdata in official statistics and there are several projects concerned with these issues coming along. In this research paper the authors will explore the current state of establishment of infrastructure for researchers' RA to statistical microdata in ESS and they will propose a model for development of infrastructure for researchers' RA to statistical microdata in Croatia with the aim of integrating the Croatian system for RA to statistical microdata in the currently developing ESS's infrastructure for the same purpose.

Key words: European Statistical System (ESS), Statistical Microdata, Remote Access (RA)

Introduction

Through history it has been a common practice for European countries to provide public only with aggregated statistical information. These aggregated data, also known as macrodata, present results from publicly funded statistical surveys. Data from which macrodata are produced, i.e. microdata, would reside in the premises, i.e. on the servers of organisations performing these surveys, or some partner archiving organisation responsible for safe preservation of statistical data. However, the prevailing trend and orientation of global politics is changing towards opening access to publicly funded data, collected in publicly

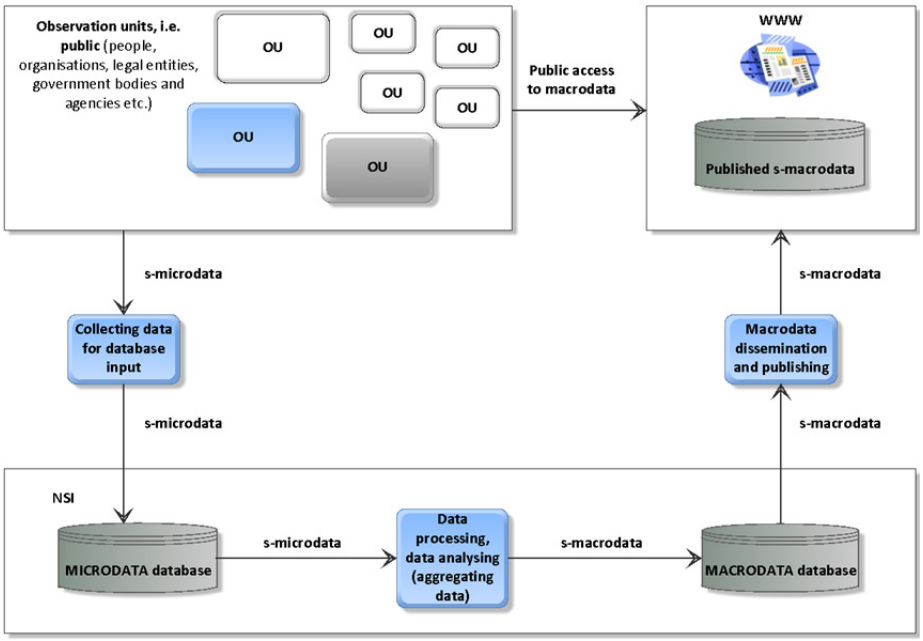
funded projects. Yet, another trend in information and communication technologies is towards turning closed organisations' information systems into open systems, available on the Internet through a set of web services, i.e. orientation towards service oriented architecture (SOA) and service applications available for global society users. The same trend is introduced in statistical information systems management. The main objective in development of information systems in statistics is set to harmonisation of statistical processes, standardisation of procedures and establishment of cooperation between different countries through means of information and communication technologies and standardised datasets. New systems, oriented towards opening access to raw data for the public, are on their way to become present in countries all over the world. These systems are considered necessary in order to have more transparent and effective society. This new approach will bring many changes and will have a big impact on everyday life in today's knowledge-driven society. Special concerns in statistics area are the legal compliance of the new social system and safe implementation of the needed services, especially due to the fact that different countries deploy different legislatives about statistical microdata confidentiality. Statistical microdata are an example of confidential datasets, so they will probably never be completely open for public use. However, they can be opened to some extent[1] for researchers' use.

Data throughout statistical surveys

To be able to provide statistical indicators about society and environment to the public, NSOs in countries collect microdata about statistical observation units (people, organisations, events, etc.). NSOs process those microdata about statistical observation units and produce aggregated data, i.e. statistical macrodata. Statistical macrodata are disseminated by various publication channels[2]. Statistical macrodata are publicly owned goods. The system presented by the model in the Picture 1 should be useful for explaining micro and macrodata-flows between public and NSO. The authors have illustrated the basic model of NSO, collecting microdata from public and in return providing public access to statistical macrodata. Amongst collecting microdata through surveys from public, there are some other possible sources of microdata for NSOs, like administrative registers, providing microdata input. Having that in mind the model presented in the Picture 1 should be perceived as a basic model of collecting statistical microdata, presenting basic methods of collecting statistical microdata, i.e. statistical surveys.

[1] Usually anonymised statistical microdata are opened for use to the research community.

[2] Usually by printed or web editions or by interactive database warehouses available on the web which enable browsing and analysing statistical macrodata.



Picture 1. National Statistical Office providing public access to macrodata (basic model of NSO)

NSOs have legal right[3] to collect confidential data about statistical survey units (persons, institutions, events, etc.) as they are objects (input data) for statistical surveys.

Microdata Confidentiality

Microdata are protected by legislatives which differ from country to country.[4] As stated in Data Without Boundaries (2012a) " (...) highly-detailed and confidential microdata (...) should be understood as data about individual statistical units that, due to the level of details provided, contain the risk of disclosure for the specific statistical unit. Therefore such data can only be accessed under technical, organizational and statistical restrictions." Definitions of confidential

[3] These rights can only be practiced as stated in legislatives covering statistical survey in question and only persons involved in the statistical survey have access to data in question. All staff members involved in data processing in NSOs and therefore having access to confidential data are legally obligated to secure and prevent data from being disclosed to other persons during and after the period of involvement in statistical processes, i.e. statistical surveys, and are especially forbidden to use statistical data for nonstatistical purposes of any kind.

[4] Data classification should be and it already is in place in many countries.

microdata records are not limited just on data presenting one statistical unit but "definitions of confidential microdata depend on national legislation, the design of the survey (sample size, included variables) and the means of access (e.g. Remote Access, Safe Centre, Download)."[5] This means that small sets of records can be considered too revealing and the possibility of disclosure of a single record in a recordset could be the reason for classifying the whole recordset as confidential.

Researchers' access to Microdata

Research communities and governmental bodies are using statistical microdata as they are useful when put to some secondary use and made available to researchers who are in turn able to produce new and meaningful representations for the society. In Data Without Boundaries (2012a) a clear conclusion is made that "only real Remote Access (RA) systems would meet the cumulative challenge of the high demand for confidential microdata, limited resources inside the NSIs (...) and researcher's needs."

ESS projects for development of RA to Microdata

As presented by the results in Data Without Boundaries (2012a) some European countries, namely Denmark, France, Slovenia, the Netherlands, Sweden and United Kingdom, offer researchers RA service to their official statistical microdata. Also, there is a Safe Centre in Eurostat[6] in Luxemburg where researchers can remotely access microdata from some statistical surveys of several EU member countries. At this moment these procedures are very complex and time consuming and there is a need for establishing a more cooperative, inter-connected and user-friendly system. Also, there are running projects, like ESSnet's "Decentralised and Remote Access to Confidential Data in the ESS" - DARA, concerned with the establishment of European Safe Centre for Access to statistical microdata in all Europe's NSOs and partner organisations, as well as with connecting Safe Centres in countries with the servers in Luxemburg. The project "Decentralised Access to EU microdata sets" ended in January 2010 and it produced a feasibility study for this establishment and was also a trigger for the DARA project. Eurostat is in the process of establishing "Secure Infrastructure for Confidential data access and sharing" (VIP-SICON) project to build a remote access system based on Citrix and the software development for this purpose has just begun, so at this moment it is still not evident how this system will operate. The Data Without Boundaries project is aimed at developing an easy and equal access rights and infrastructure to official microdata for researchers

[5] Data Without Boundaries (2012a).

[6] Eurostat is the Statistical Office of the European Union, i.e. central statistical authority in the ESS.

for the European Research Area through coordination of existing infrastructures, Council of European Social Science Data Archives (CESSDA[7]) and the ESS.

Available RA Centres in Europe

A definition of Remote Access to microdata, as presented in Data Without Boundaries (2012a), is „any kind of terminal or desktop solution that uses a secured connection to the servers of the respective data provider, whereby the user can see the real, highly detailed and confidential microdata.“ The first country to offer their researchers RA to microdata was Denmark in 2002, followed by UK and Slovenia in 2004 and Sweden in 2005. In 2006 the Netherlands joined this community and two more countries in 2009, namely France and again UK8 introduced microdata RA, followed by Germany in 2011. Other countries still have to develop similar systems based on existing models with the aim of being compatible to others, as the movement towards cooperation in statistical area is evident and expected by authorities like EU Commission9 and Eurostat.

Common and different RA areas in available centres

Although, at this time, there is no case of two exactly the same RA centres in Europe – they all, according to the study performed and results presented in Data Without Boundaries (2012a), have some common areas in the way they are assembled and organised to work. Firstly, in all RA centres confidential data does not leave the research data centre. Users basically connect to RDC's[10] server with thin clients[11] or via regular PCs, which are only able to trigger execution of commands on the protected RDC's server, where data is processed. Secondly, after being processed on RDC's server, data are sent – not directly to the user's PC but – to an RDC location. This location is usually user's outbox folder in RDC where the data stays until examined by an NSO's staff member. The data can be sent to a researcher only if approved to be safe for sending. All data transmitted to users are kept for some period in the archive of the RDC (3 months or so). At all the RDC's there is a process of accreditation before a re-

[7] CESSDA at the moment has 20 member countries in Europe.

[8] Different service then before in 2004.

[9] EU Commission, along with Eurostat, has a big impact on the development of official statistical systems in Europe. The two organisations are active in introducing standards and establishing harmonised statistical systems in Europe.

[10] RDC – Remote Data Centre.

[11] "A thin client is desktop computer that is only used to provide input and output to the user. The calculations are done on a server. According to that, thin clients are equipped with a limited operating system that allows only restricted use of functionalities (Data Without Boundaries (2012a))."

searcher can get access. For this purpose passwords are used, given IP-ranges, smart cards with fingerprints, i.e. biometric technologies[12], smart cards that create one time password (OTP), hardware components[13] or token with RSA SecureID. Most of the RDCs use more than just one of these methods for accreditation. All RDCs have sophisticated user account management and use Microsoft active directory for user management. The internet (or private government network[14] enables remote connection between researchers and data, with relatively standardised technologies for encryption of communication. Some of the applied encryption methods are SSTP, SSl-VPN, Cisco-tunnel, RDP (Microsoft) and ICA Basic encryption (Citrix). Surveillance is carried out during research activities (methods like monitoring sessions, logs, log-on and log-off information etc.). At RDC's standard statistical software is usually made available free of charge (SAS, STATA, SPSS). Additional software is possible, but users have to cover the cost.

Building an integrated RA Centre for Microdata in Europe

Combined statistical microdata from various Data Centres available across Europe using RA services could be a precious source for researchers conducting comparisons and drawing important conclusions throughout the Europe, thus enabling better informed political decisions for society and economy. All this potential should not be neglected, especially in today's knowledge-driven economy. Many people agree on the notion that the so called "online borders" between European countries should be removed, thus enabling cooperation and even global proactive involvement of society, using global microdata sets from different countries. It is obvious, however, that these changes should be introduced with careful and reasonable consideration of possible benefits and possible threats for the society in the long term, and especially careful planning of new legal framework. Yet, at this moment researchers are facing complex procedures and have to be prepared to invest a lot of time and effort when in need of statistical microdata as they are available at the safe centre in Eurostat, Luxemburg with the exception of countries enabling this kind of remote access locally[15] to their citizens, and yet another different set of rules is in place for the "noncitizens". The process of getting the authorisation privileges is rather complex and it takes some time before users can use the data. European Commis-

[12] Biometrics is being used currently in France.

[13] In France the so called SD-box-hardware certificate device.

[14] Private government network for researchers' RA to statistical microdata is used in UK only at this moment.

[15] These countries are discussed in this paper, but they only provide access for their citizens, with an exception of Germany, which also provides remote access from US to the data in Germany, for a few US citizens involved in projects at IAB in Germany.

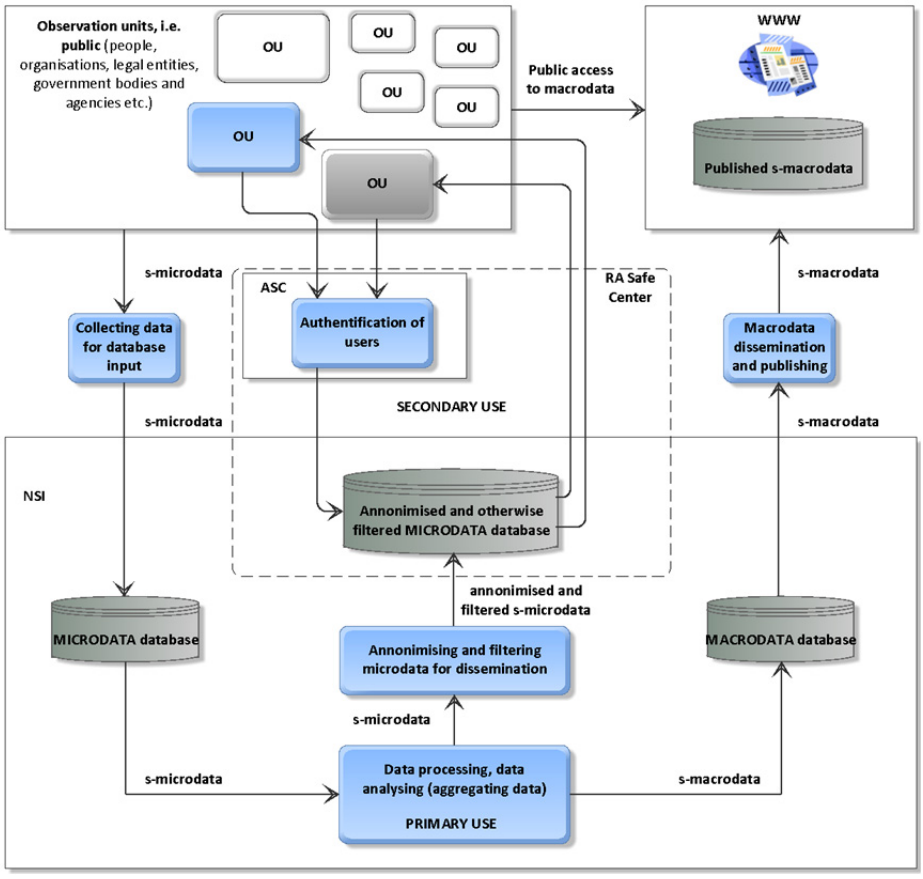
sion is on its way to simplify all these processes by implementing international e-signatures and international e-certificates on the Internet. It is presumable to expect their integration into governmental services for public over the Internet. There are some projects coming along at this moment with the aim of changing the state of RA to microdata in Europe. ESSnet projects deal with the problem of establishing an RA to microdata in Eurostat Safe Centre from local NSOs in European countries. DARA will deliver a running system connecting Safe Centres from countries in Europe to microdata at the Eurostat in Luxemburg. It already had two successful connections established from two countries in Europe until now. The project is planned to end in October 2013.

Proposal of the model for microdata RA in Croatia

The main statistical body of Croatia is Central Bureau of Statistics (CBS). Since Croatia became EU member in the July 2013 Croatia also took new obligations to fulfil requirements of the EU, Eurostat and other organisations involved in the ESS establishment. At the moment of writing this paper, Croatia is not a member of the CESSDA association. Following the global trend toward open government initiatives Croatian government has signed Open Government Partnership agreement in August 2011. Open Government Partnership, as stated in their web release, is a new multilateral initiative that aims to secure concrete commitments from governments to promote transparency, empower citizens, fight corruption, and harness new technologies to strengthen governance.[16] Recently signed G8 Open Data Charter[17] states that "access to data allows individuals and organisations to develop new insights and innovations that can improve the lives of others and help to improve the flow of information within and between countries. (...) Today, people expect to be able to access information and services electronically when and how they want. Increasingly, this is true of government data as well. (...) Open data can increase transparency about what government and business are doing." Further in the Charter the G8 countries state that they will "establish an expectation that all government data be published openly by default (...) while recognising that there are legitimate reasons why some data cannot be released; (...) release high-quality open data that are timely, comprehensive, and accurate. To the extent possible, data will be in their original, unmodified form and at the finest level of granularity available; (...) make sure that data are fully described, so that consumers have sufficient information to understand their strengths, weaknesses, analytical limitations, and security requirements, as well as how to process the data; and release data as early as possible." Also, they recognise "statistics (National Statistics, Census, infrastructure, wealth, skills)" as "areas of high value, both for improving our democracies and encouraging innovative re-use of data."

[16] Open Government Partnership, August 2011, http://www.opengovpartnership.org/about.

[17] G8 Open Data Charter, UK Cabinet Office, 2013.



Picture 2. National Statistical Office providing access to macrodata for public users and anonymised and filtered microdata for researchers’ RA and secondary analysis purposes

Remote Access to statistical microdata is one way to accomplish these goals. Croatia is an example of country without RA service centre for microdata in statistics. It is reasonable to expect that Croatia will be involved in developing Safe Centres for RA to microdata in the country connecting them with other Safe Centres in the ESS in the near future. Having that in mind, in the Picture 2 the authors are presenting a model for microdata RA Safe Centre for Croatia. In the proposed model shaded observation units present a researcher and a governmental institution employee seeking access to microdata. As illustrated, these users are able to get access to statistical microdata after passing through security checking by Authentication Service Centre (ASC). Considering the complex procedures for data classification in official statistics, which vary from

one statistical domain to the other, it is reasonable to expect a considerable effort will be invested in preparing automated, i.e. computerised, procedures for data classification in RA systems in order to have a fully automated system in place. For RA to microdata special authentication services must be put in place. Croatia has already established a similar service, called AAI identity, at the Croatian Academic and Research Network (CARNet). AAI provides electronic identity authentication service for members of the research community. However, Open Data initiatives are aimed at opening publicly funded data to all public members, not just to the researchers and government officials. If those systems are to be introduced and made available through a set of online web services, it is reasonable to expect that in the end these systems will be very complex, especially in the security layers implemented and authentication services used for various user profiles. Certainly, they could benefit from a planned introduction of NIAS[18] in Croatia and this system should be developed with the aim of being compatible with various systems, hopefully it will be useful for implementing in researchers' authentication for statistical Microdata RA. It would be especially wise to think beyond the borders of Croatia. In that sense Croatia, as a member of ESS, should join the combined network for establishment of RA to statistical microdata in the ESS.

Conclusion

The full potential of statistical data is not derived by using aggregated data, i.e. macrodata. It should be possible to move beyond and to be able to connect, at least research community members, to safely protected statistical microdata using RA services and enabling easy access. Using that data, which possibly hold many undiscovered secrets and relationships, could introduce lots of benefits for society and empower individuals and organisations. In the last ten years outstanding progress has emerged in some European countries, and different countries have developed and implemented different sets of network infrastructure, protocols and ways of researchers' RA to statistical microdata implementing various ways of authentication of users and providing security for accessible statistical microdata. It is reasonable to expect future involvement of Croatia in this area as Croatia has just became a member of the EU and a full member of ESS. We have proposed the model for development of Safe Centre for RA to statistical microdata in Croatia, critically stressed the importance of harmonising applied procedures and careful introduction and development of system for microdata RA in Croatia with the aim of integrating this system into the ESS's system for RA to statistical microdata.

[18] NIAS – National Identification and Authentification System.

References

Brandt, Maurice. Decentralised and Remote Access to Confidential Data in the ESS (ESSnet DARA), Overview and State of the Art. 2012. http://www.safe-centre.info/wp-content/uploads/2012/07/ESSnet_DARA_Overview_Maurice_Brandt_20120328.pdf (05/02/2013)

Brandt, Maurice; Zwick, Markus. Improvement of data access – The long way to remote data access in Germany. 2011. http://www.forschungsdatenzentrum.de/publikationen/veroeffentlichungen/fdz_arbeitspapier-39.pdf (05/02/2013)

Council of European Social Science Data Archives (CESSDA). 2013. http://www.cessda.org/about/ (05/02/2013)

Data Service infrastructure for the Social Sciences and Humanities. About Dasish. 2013. http://dasish.eu/about_dasish/ (05/02/2013)

Data Without Boundaries. Deliverable D4.1. Report on the current state of the art of current SC in Europe. 2012a. http://www.dwbproject.org/export/sites/default/about/public_deliveraples/d4_1_current_sc_in_europe_report_full.pdf (05/02/2013)

Data Without Boundaries. Deliverable D5.1. Report on concept for and components of European Service Centre for official statistics, 2012b. http://www.dwbproject.org/export/sites/default/about/public_deliveraples/d5_1_european_service_centre_report.pdf (05/02/2013)

Desai, Tanvi; Ritchie, Felix. Effective Researcher Management. 2009. http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2009/wp.15.e.pdf (05/02/2013)

Desai, Tanvi. ONS Virtual Microdata Laboratory: Strategic Recommendations 2011-2015. 2011. http://www.esrc.ac.uk/search/search-page.aspx?q=*%3A*&current_page=1&items_per_page=10&sort_order=relevance&filters=on&author=Tanvi%20Desai&tab=outputs (05/02/2013)

ESSnet Home of DARA. ESSnet on Decentralised and Remote Access to Confidential Data. 2013. http://www.safe-centre.info/ (05/02/2013)

European Commission. Commission Regulation (EU) No 520/2010. 2010. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:151:0014:0015:EN:PDF

European Commission. Commission Regulation (EU) No 557/2013. 2013. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:164:0016:0019:EN:PDF

European Commission. Digital Agenda for Europe, A Europe 2020 Initiative. 2010. http://ec.europa.eu/digital-agenda/digital-agenda-europe (05/02/2013)

European Strategy Forum on Research Infrastructures (ESFRI). Strategy report on Research Infrastructures, Roadmap 2010. 2010. http://ec.europa.eu/research/infrastructures/pdf/esfri-strategy_report_and_roadmap.pdf#view=fit&pagemode=none (05/02/2013)

G8 Open Data Charter, UK Cabinet Office, 2013, https://www.gov.uk/government/publications/open-data-charter

Horton, Laurence; Katsanidou, Alexia. Purposing your survey: archives as a market regulator, or how can archives connect supply and demand?. 2011. http://iassistdata.org/downloads/iqvol35_horton.pdf (05/02/2013)

Manifesto The New Global Data Generation. Simplifying and Guaranteeing Access and Sharing in e-Science Scenarios. 2012. http://dasish.eu/manifesto/Manifesto2012-03-14.pdf/ (05/02/2013)

Open Government Partnership. Commitment of Croatia. 2011. http://www.opengovpartnership.org/countries/croatia (05/02/2013)

Ritchie, Felix. Secure Access to confidential microdata: four years of the Virtual Microdata Laboratory. 2008. http://www.ons.gov.uk/ons/rel/elmr/economic-and-labour-market-review/no--5--may-2008/secure-access-to-confidential-microdata--four-years-of-the-virtual-micro data-laboratory.pdf (05/02/2013)

Vrana, Radovan. Supporting e-Science: Scientific Research Data Curation. 2011. http://infoz.ffzg.hr/INFuture/2011/ConferenceProgram.aspx (05/02/2013)

Long-term Preservation of Validity of Electronically Signed Records

Hrvoje Brzica*
Boris Herceg*
Financial Agency – FINA
Ulica grada Vukovara 70, 10000 Zagreb, Croatia
Hrvoje.Brzica@fina.hr, Boris.Herceg@fina.hr

Hrvoje Stančić
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
hstancic@ffzg.hr

Summary

The authors explain the context in which electronic records are being preserved. They explain the concept of authentic electronic records and proceed with the analysis of the technologies supporting trust in electronic records. They start by explaining the Public Key Infrastructure as the requirement for electronic signatures, digital certificates, the concept of non-repudiation, trusted archive service, timestamps and trusted digital timestamping. Further, they analyse formats of electronic signatures – XMLDSig, XAdES, CAdES, PAdES – and their possible influence on the long-term preservation of validity of electronically signed records. The authors conclude that although strict requirements of certain types of electronic signatures can ensure authenticity, integrity and non-repudiation of preserved records, they will still require preservation action on the level of medium and files.

Key words: electronic signature, digital certificate, non-repudiation, trusted archive service, timestamp, XMLDSig, XAdES, CAdES, PAdES

Introduction

Digital documents are being created in every segment of modern business. Those documents may or may not be part of a document management system. Increasingly they are not printed and signed but are digitally signed instead. At that moment they become records. They are used in the process of business and later, when they lose their immediate business value, are archived and may or

* The positions put forward in this article are solely those of the authors' and do not necessarily reflect the positions of FINA.

may not become part of a records management system and stored within a digital archive solution. If those records are to be preserved in the long-term their characteristics of authenticity, reliability, integrity and usability[1] have to be preserved.

"Authenticity is not a single concept, but involves different aspects that can be associated with an object:

- A traceable path from the object's original to its current ownership.
- Measures and techniques for safeguarding against and/or recognizing modifications.
- Techniques for establishing the use of original materials.

Usage and context define how these aspects are defined for individual classes of objects."[2] Therefore, as Duranti says, "a document is authentic if it can be demonstrated that it is precisely as it was when first transmitted or set aside for preservation, and if its reliability, i.e., the trustworthiness it had at that moment, has been maintained intact."[3]

This shows that the concept of long-term preservation of digital records with the abovementioned characteristics requires a complex digital solution. The aim of this paper is to give a context in which modern digital documents and records are being created, to explain the technologies required to support that process and to analyse their influence on long-term preservation of validity of electronically signed records. Therefore, the concepts of electronic signatures, digital certificates, non-repudiation, trusted archive service, timestamps and trusted digital timestamping will be explained. For better understanding of those concepts, the concept of Public Key Infrastructure needs to be shortly explained.

"Public Key Infrastructure (PKI) represents complex information infrastructure, which is used to manage electronic identities. Basis of PKI relies on asymmetric encryption. Asymmetric encryption actually relies on mathematically related key pair, one called the public key, and another called private key, generated to be used together. (...) The private key is kept secret and used only by its owner, while public key is made available to anyone who wants it."[4] Modern systems can easily use the keys with length of 2048 characters which are impossible to break even by todays supercomputers.

[1] As defined by ISO 15489: Information and documentation – Records management, 2001.

[2] Van Diessen, Raymond J. and van der Werf-Davelaar, Titia, *Authenticity in a Digital Environment*, IBM / Koninklijke Bibliotheek Long-Term Preservation Study Report Series, No. 2, IBM Netherlands, Amsterdam, December 2002, p. 3, http://www.kb.nl/sites/default/files/docs/2-authenticity.pdf (22.2.2013)

[3] Duranti, Luciana, The Concept of Electronic Record, in: Duranti, Luciana, Eastwood, Terry and MacNeil Heather, *Preservation of Integrity of Electronic Records*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002, p. 27.

[4] Jacobs, J., Clemmer, L., Dalton, M., Rogers, R., Posluns, J., *SSCP Study Guide*, Syngress Publishing, 2003, pp. 330-331.

Technologies and concepts supporting trust in electronic records

Electronic signature

There are two types of electronic signatures – basic (usually referred to just as "electronic signature") and advanced. The European Telecommunications Standards Institute (ETSI) defines electronic signature as "essentially the equivalent of a hand-written signature, with data in electronic form being attached to other electronic subject data (invoice, payment slip, contract, etc.) as a means of authentication. Electronic signature is not just a 'picture' of the hand written signature. It is a digital signature that uses a cryptographic transformation of the data to allow the recipient of the data to prove the origin and integrity of the subject data."[5] European legislation in the Directive 1999/93/EC states that an electronic signature needs to meet the following requirements in order to become an advanced electronic signature[6]:

a) it is uniquely linked to the signatory[7];
b) it is capable of identifying the signatory;
c) it is created using means that the signatory can maintain under his sole control; and
d) it is linked to the data to which it relates in such a manner that any subsequent change of the data is detectable.

Digital certificates

Digital certificates are digital records used for confirming the identity of a person, an organisation or a machine. Digital certificate is valid during certain period of time and contains several additional elements. The Directive 1999/93/EC allows issuing of the so called *qualified certificate* which is based on the RFC 3039 standard and implements the concept of non-repudiation. The qualified certificate must in particular include[8]:

a) an indication that the certificate is issued as a qualified certificate;
b) the identification of the certification-service-provider and the State in which it is established;
c) the name of the signatory or a pseudonym, which shall be identified as such;

[5] Electronic signature, ETSI, 2012, http://www.etsi.org/index.php/technologies-clusters/ technologies/security/electronic-signature (9.7.2013)

[6] Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures, Official Journal L 013, 19 January 2000, pp. 12-20, http://europa.eu/legislation_summaries/information_society/other_policies/l24118_en.htm (11. 7.2013)

[7] "A person who holds a signature-creation device and acts either on his own behalf or on behalf of the natural or legal person or entity he represents." (Directive 1999/93/EC)

[8] Directive 1999/93/EC.

d) provision for a specific attribute of the signatory to be included if relevant, depending on the purpose for which the certificate is intended;
e) signature-verification data which correspond to signature-creation data under the control of the signatory;
f) an indication of the beginning and end of the period of validity of the certificate;
g) the identity code of the certificate;
h) the advanced electronic signature of the certification-service-provider issuing it;
i) limitations on the scope of use of the certificate, if applicable; and
j) limits on the value of transactions for which the certificate can be used, if applicable.

Certification authority and registration authority
In the PKI infrastructure digital certificates are issued and revoked by the Certification Authority (CA). CA is organised as a hierarchy within which a Root CA is used as the highest entity, trusting itself, while all other, hierarchically lower entities trust the Root CA. The idea is that every identified entity receives digital signature, i.e. a certificate of its public key which, in turn, can be used for confirm its identity. The procedure is that CA uses its private key to sign the digital certificate of an entity, and the identity of that entity can be checked by using CA's public key. It is important to mention that, upon the request for certification by an entity, "a CA checks with a Registration Authority (RA)[9] to verify information provided by the requestor of a digital certificate. If the RA verifies the requestor's information, the CA can then issue a certificate."[10]

Non-repudiation
Non-repudiation is a characteristic of a record that prevents any signatory to deny the action taken or the content of a record. In the Croatian legislation non-repudiation is associated with the advanced digital signature which is based upon qualified certificate. For a record to achieve and preserve characteristic of non-repudiation it is necessary to ensure:

1. digital identity of signatories,
2. real-time revocation of digital signature rights,

[9] Registration authority, as a part of PKI infrastructure, is "an authority in a network that verifies user requests for a digital certificate and tells the certificate authority (CA) to issue it." Search Security, s.n. registration authority, January 2006, http://searchsecurity.techtarget.com/definition/registration-authority (9.7.2013)

[10] Search Security, s.n. certificate authority (CA), June 2007, http://searchsecurity.techtarget.com/definition/certificate-authority (9.7.2013)

3. time-stamping of digital signatures after checking the list of revoked certificates, which ensures the validity of electronic signature at the time of signing, and
4. secure long-term preservation.

Trusted archive service

Dumortier and Eynde explain that a trusted archive service (TAS) "should maintain a set of applications (viewers as well as signature validation applications) together with the corresponding platforms (hardware, operating systems) or at least an emulator of such applications and/or environment in order to guarantee that the signature of the document can still be validated years later. To achieve this goal in the best possible way, the TAS must only accept documents in a format that can still be understood when the format will no longer be in use. Only open file formats that are vendor-independent qualify for long term archiving. (...) Every TAS must therefore publish a list of supported document formats. Such a list may be exhaustive or very restricted. Every time a document is submitted, the TAS must check the format before accepting it for archiving."[11]

Timestamp and trusted digital timestamping

According to Wallace et al. a digital timestamp is an attestation generated by a Time Stamping Authority (TSA) – a trusted service – that a data item existed at a certain time.[12] Ćosić and Bača explain that "time stamps are typically used for logging events, in which case each event in a log is marked with a time stamp. In file systems, time stamp may refer to the stored date/time of the file creation or modification. *Trusted time stamping* is the process of securely keeping track of the creation and modification time of a document. (...) Trusted TSA can be used to prove the consistency and integrity of digital evidence in every stage of its existence."[13]

Formats of electronic signatures

In the analysis so far the technologies and concepts supporting trust in electronic records were explained. It was shown that the concept of electronic signature can be viewed as the basis for developing all other technologies. Further,

[11] Dumortier, Jos and Eynde, Sofie Van den, Electronic Signatures and Trusted Archival Services, DAVID Project (2000-2003), p. 7, http://www.expertisecentrumdavid.be/davidproject/teksten/DAVIDbijdragen/Tas.pdf (8.7.2013)

[12] Wallace, C., Pordesch, U. and Brandner R., Long-Term Archive Service Requirements, IETF Trust, 2007, p. 5, http://tools.ietf.org/pdf/rfc4810.pdf (9.7.2013)

[13] Ćosić, Jasmin and Bača, Miroslav, (Im)Proving Chain of Custody and Digital Evidence Integrity with Time Stamp, MIPRO – Proceedings of the 33rd International Convention, 2010, pp. 1227-1228, http://czb.foi.hr/upload/datoteke/10_400%281%29.pdf (9.7.2013)

the realisation of digital signatures will be explained through analysis of the most important formats of electronic signatures – XMLDSig, XAdES, CAdES and PAdES.

XMLDSig

XML (Extensible Markup Language) Signature is defined by the W3C Recommendation[14]. In the literature it is referred to as XMLDSig, XML-DSig or XML-Sig. The W3C Recommendation states that "XML Signatures can be applied to any digital content (data object), including XML. An XML Signature may be applied to the content of one or more resources." One can differentiate between[15]:

- detached signature – an XML signature used to sign a resource outside its containing XML document, i.e. the signature is over content external to the `Signature` element;
- enveloped signature – signature is child; it is used to sign some part of its containing document, i.e. the signature is over content found within an `Object` element of the signature itself:
- enveloping signature – signature is parent; it contains the signed data within itself, i.e. the signature is over the XML content that contains the signature as an element. The content provides the root XML document element. Obviously, enveloped signatures must take care not to include their own value in the calculation of the `SignatureValue`.

"XML Signatures are applied to arbitrary digital content (data objects) via an indirection. Data objects are digested, the resulting value is placed in an element (with other information) and that element is then digested and cryptographically signed."[16]

XAdES

XAdES (XML Advanced Electronic Signature) "extends XMLDSig specification into the domain of non-repudiation by defining XML formats for advanced electronic signatures that remain valid over long periods and are compliant with the European 'Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures' and incorporate additional useful information in common uses cases. This includes evidence as to its validity even if the signer or verifying party

[14] XML Signature Syntax and Processing (Second Edition), W3C Recommendation, The Internet Society & W3C, 10 June 2008, http://www.w3.org/TR/xmldsig-core/ (10.7.2013)

[15] XML Signature, Wikipedia, June 2013, http://en.wikipedia.org/wiki/XML_Signature (10.7. 2013).
XML Signature Syntax and Processing (Second Edition), W3C Recommendation.

[16] XML Signature Syntax and Processing (Second Edition), W3C Recommendation.

later attempts to deny (repudiates) the validity of the signature. An advanced electronic signature aligned with the present document (i.e. XAdES specification) can, in consequence, be used for arbitration in case of a dispute between the signer and verifier, which may occur at some later time, even years later."[17] In relation to XMLDSig, XAdES specification adds six additional, mutually nested, forms (see Figure 1):

1. XAdES (also referred to as XAdES-BES – basic electronic signature) – basic form defining elements for authentication and protection of integrity of records but not providing non-repudiation of its existence.
2. XAdES-T (Timestamp) – addition of the timestamp ensures non-repudiation.
3. XAdES-C (Complete validation data) – builds up on the XAdES-T by adding references to the set of data supporting the validation of the electronic signature (i.e. the references to the certification path and its associated revocation status information). This form is useful for those situations where such information is archived by an external source, like a trusted service provider.
4. XAdES-X (eXtended validation data) – builds up on XAdES-C by adding timestamps to protect against the risk that any keys used in the certificate chain or in the revocation status information may be compromised.
5. XAdES-X-L (eXtended validation data incorporated for the Long term) – builds up on XAdES-X by adding the validation data (i.e. certificates and revocation values) for those situations where the validation data are not stored elsewhere for the long-term.
6. XAdES-A (Archiving validation data) – builds up on XAdES-X-L by adding time-stamps for archiving signatures

CAdES

CAdES (CMS Advanced Electronic Signatures) is a set of extensions to CMS (Cryptographic Message Syntax) signed data. It "defines a number of electronic signature formats, including electronic signatures that can remain valid over long periods. This includes evidence as to its validity even if the signer or verifying party later attempts to deny (repudiates) the validity of the electronic signature."[18] Similar to XAdES, CAdES specification defines six profiles, each building up on the previous one: CAdES – basic form, CAdES-T (Timestamp), CAdES-C (Complete), CAdES-X (eXtended), CAdES-X-L (eXtended Long-term) and CAdES-A (Archival). The difference between the two specifications

[17] XML Advanced Electronic Signatures (XAdES), W3C Note 20 February 2003, ETSI, 2003, http://www.w3.org/TR/XAdES/ (10.7.2013)

[18] CMS Advanced Electronic Signatures (CAdES) Technical Specification, ETSI TS 101 733 v1.7.4, July 2008, p. 8, http://www.etsi.org/deliver/etsi_ts/101700_101799/101733/01.07.04_60/ts_101733v010704p.pdf (10.7.2013)

is that while CAdES renders signature as binary data, XAdES provides an XML solution.

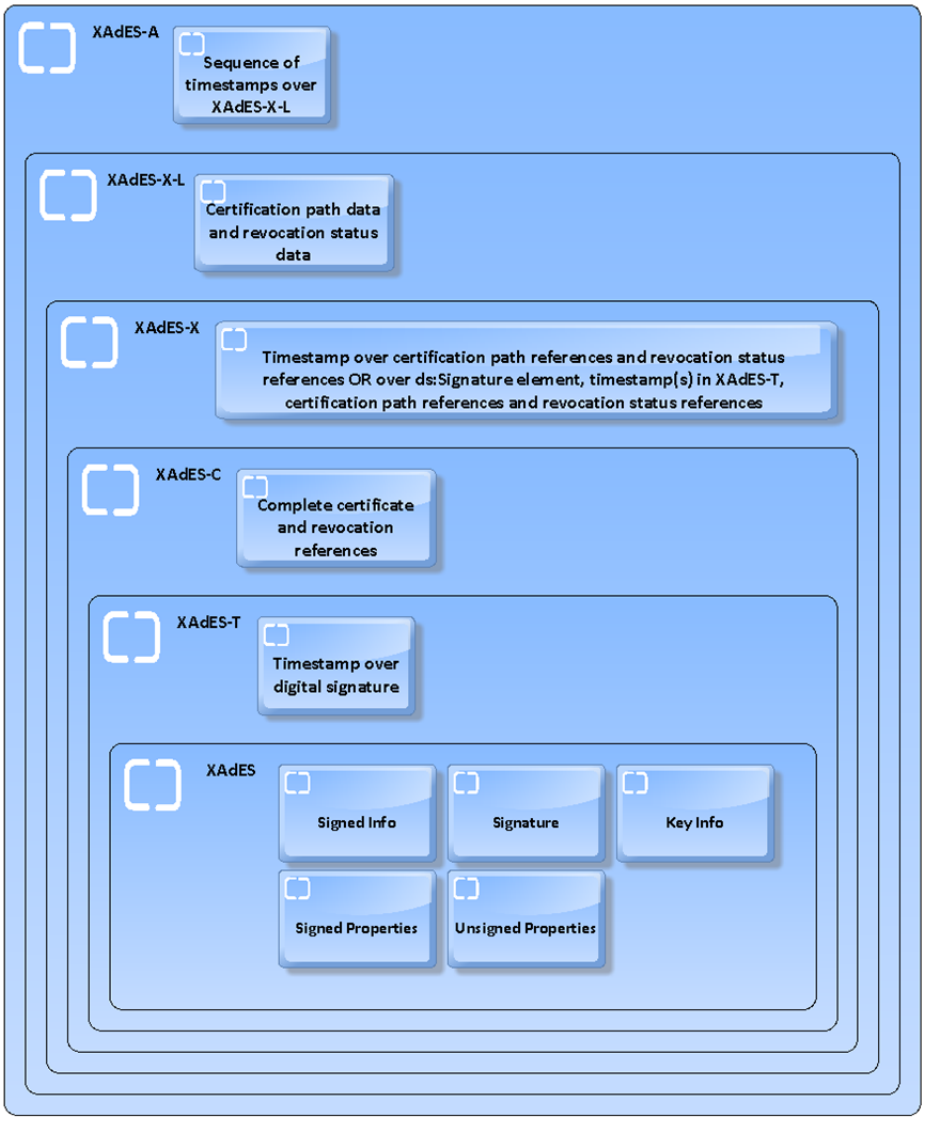


Figure 1: Structure of XAdES specification forms

PAdES

PAdES (PDF Advanced Electronic Signature) "articulates the same capabilities featured in CAdES and XAdES for PDF. (...) PAdES differs from CAdES and XAdES in that it applies only to PDF documents and defines requirements that PDF viewing and editing software must follow when using digital signatures in

PDF documents. As the standard for viewable documents, PDF also defines how a signature can be displayed as it might with an ink-on-paper signature at a particular position on a particular page, and how digital signatures can be integrated with the form-filling features of PDF. This is a key factor that distinguishes it from CAdES and XAdES, which are more suited for applications that may not involve human-readable documents."[19] PAdES specification is realised in 6 parts:

Part 1 – PAdES Overview – a framework document for PAdES[20]

Part 2 – PAdES Basic – Profile based on ISO 32000-1: specifies a PDF signature as specified in ISO 32000-1:2008 that enables greater interoperability for PDF signatures by providing additional restrictions beyond those of ISO 32000-1.[21]

Part 3 – PAdES Enhanced: incorporates the signature time-stamp attribute as optional making the signature effectively a CAdES-T form.[22]

Part 4 – PAdES Long Term: uses an extension to ISO 32000-1 called Document Security Store (DSS) to carry such validation data as necessary to validate a signature, optionally with Validation Related Information (VRI) which relates the validation data to a specific signature.[23]

Part 5 – PAdES for XML Content: profile for usage of arbitrary signed (with XAdES signatures) XML document that is embedded within a PDF file, for providing integrity, authentication and non-repudiation services on the data objects that are signed with the XAdES signature.[24]

Part 6 – Visual Representations of Electronic Signatures: specifies requirements and recommendations for the visual representations of advanced electronic signatures (AdES) in PDFs[25]. The signature appearance is created by the signer and any identification included in the signature appearance is not directly verifiable by the AdES signature. However, this information may be visually checked against the visual representation of the electronic signature (AdES) verification.[26]

[19] The AdES family of standards: CAdES, XAdES, and PAdES. Implementation guidance for using electronic signatures in the European Union, White paper, Adobe Systems, 2009, p. 5, http://blogs.adobe.com/security/91014620_eusig_wp_ue.pdf (10.7.2013)

[20] PAdES. Part 1.

[21] PAdES. Part 2, p. 8.

[22] PAdES. Part 3, p. 7.

[23] PAdES. Part 4, p. 8.

[24] PAdES. Part 5, p. 9.

[25] PAdES. Part 6, p. 5.

[26] Ibid., p. 7.

ETSI is developing Part 7 of the PAdES specification: Baseline Profile which will address e-Invoicing. The profile will identify a common set of options that are appropriate for maximizing interoperability between PAdES signatures. At the moment of writing this paper, Part 7 was still in the form of early draft.

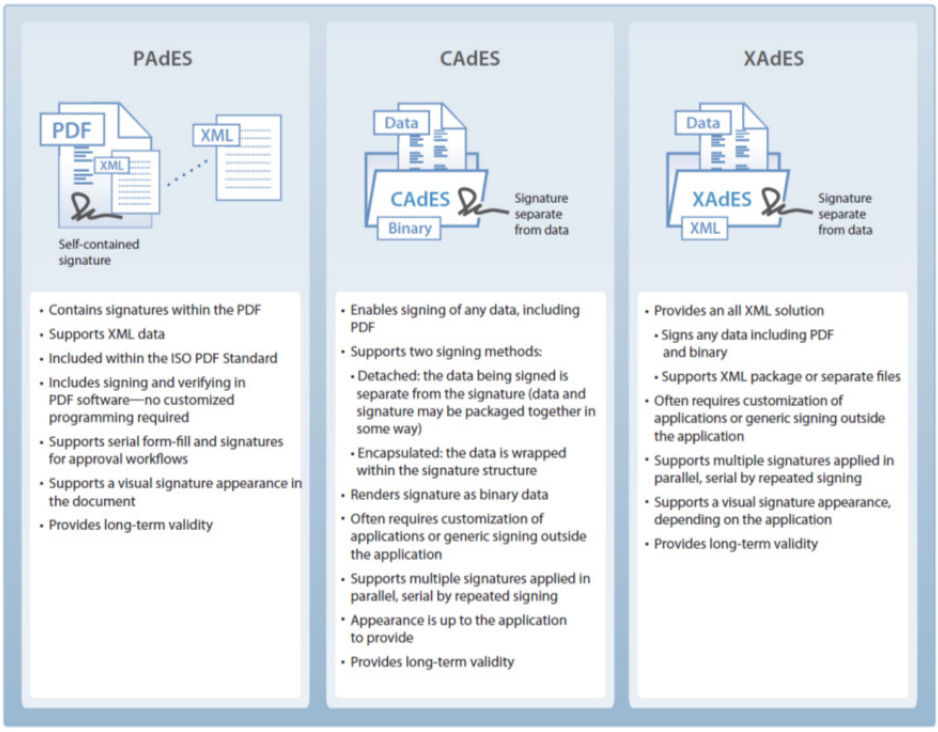


Figure 2: Comparison of PAdES, CAdES, and XAdES[27]

Conclusion

The analysis of technologies and concepts supporting trust in electronic records as well as the formats of digital signatures showed important issues that should be addressed when long-term preservation of authentic electronic records is considered. Setting up a digital archive in accordance with the discussed Directive 1999/93/EC and European Telecommunications Standards Institute's advance digital signature specifications can ensure authenticity, integrity and non-repudiation of preserved records signed by advance digital signatures because they enable possibility of checking the validation chain during the long-term preservation. However, it is important to keep in mind that majority of solutions explained in this article try to tackle only the problem of long-term preservation of electronic signatures, not the actual documents that are signed.

[27] The AdES family of standards, p. 7.

Therefore, long-term preservation of electronically signed records will, in time, still require some kind of preservation action, e.g. emulation or migration. Ideally, establishing a trusted archive service based both on advance digital signature specifications and proactive approach to digital preservation should prove the best solution.

References

Boudrez, Filip, Digital Signatures and Electronic Records, Archival Science 7(2), 2007, pp. 179-193., http://www.edavid.be/docs/digitalsignatures.pdf (12.7.2013)

CMS Advanced Electronic Signatures (CAdES) Technical Specification, ETSI TS 101 733 v1.7.4, July 2008, http://www.etsi.org/deliver/etsi_ts/101700_101799/101733/01.07.04_60/ts_101733v010704p.pdf (10.7.2013)

Ćosić, Jasmin and Bača, Miroslav, (Im)Proving Chain of Custody and Digital Evidence Integrity with Time Stamp, MIPRO - Proceedings of the 33rd International Convention, 2010, pp. 1226-1230, http://czb.foi.hr/upload/datoteke/10_400%281%29.pdf (9.7.2013)

Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures, Official Journal L 013, 19 January 2000, pp. 12-20, http://europa.eu/legislation_summaries/information_society/other_policies/l24118_en.htm (11.7.2013)

Dumortier, Jos and Eynde, Sofie Van den, Electronic Signatures and Trusted Archival Services, DAVID Project (2000-2003), pp. 1-9, http://www.expertisecentrumdavid.be/davidproject/teksten/DAVIDbijdragen/Tas.pdf (8.7.2013)

Duranti, Luciana, The Concept of Electronic Record, in: Duranti, Luciana, Eastwood, Terry and MacNeil Heather, Preservation of Integrity of Electronic Records, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002

Electronic signature, ETSI, 2012, http://www.etsi.org/index.php/technologies-clusters/technologies/security/electronic-signature (9.7.2013)

ISO 15489: Information and documentation – Records management, 2001.

Jacobs, J., Clemmer, L., Dalton, M., Rogers, R., Posluns, J., *SSCP Study Guide*, Syngress Publishing, 2003

PAdES – PDF Advanced Electronic Signature Profiles. Part 1: PAdES Overview - a framework document for PAdES, Technical Specification, ETSI TS 102 778-1 V1.1.1, July 2009, http://www.etsi.org/deliver/etsi_ts/102700_102799/10277801/01.01.01_60/ts_10277801v010101p.pdf (10.7.2013)

PAdES. Part 2: PAdES Basic - Profile based on ISO 32000-1, Technical Specification, ETSI TS 102 778-2 V1.2.1, July 2009, http://www.etsi.org/deliver/etsi_ts/102700_102799/10277802/01.02.01_60/ts_10277802v010201p.pdf (10.7.2013)

PAdES. Part 3: PAdES Enhanced - PAdES-BES and PAdES-EPES Profiles, Technical Specification, ETSI TS 102 778-3 V1.2.1, July 2010, http://www.etsi.org/deliver/etsi_ts/102700_102799/10277803/01.02.01_60/ts_10277803v010201p.pdf (10.7.2013)

PAdES. Part 4: PAdES Long Term - PAdES-LTV Profile, Technical Specification, ETSI TS 102 778-4 V1.1.2, December 2009, http://www.etsi.org/deliver/etsi_ts/102700_102799/10277804/01.01.02_60/ts_10277804v010102p.pdf (10.7.2013)

PAdES. Part 5: PAdES for XML Content - Profiles for XAdES signatures, Technical Specification, ETSI TS 102 778-5 V1.1.2, December 2009, http://www.etsi.org/deliver/etsi_ts/102700_102799/10277805/01.01.02_60/ts_10277805v010102p.pdf (10.7.2013)

PAdES. Part 6: Visual Representations of Electronic Signatures, Technical Specification, ETSI TS 102 778-6 V1.1.1, July 2010, http://www.etsi.org/deliver/etsi_ts/102700_102799/10277806/01.01.01_60/ts_10277806v010101p.pdf (10.7.2013)

Search Security, s.n. certificate authority (CA), June 2007, http://searchsecurity.techtarget.com/definition/certificate-authority (9.7.2013)

Search Security, s.n. registration authority, January 2006, http://searchsecurity.techtarget.com/definition/registration-authority (9.7.2013)

The AdES family of standards: CAdES, XAdES, and PAdES. Implementation guidance for using electronic signatures in the European Union, White paper, Adobe Systems, 2009, http://blogs.adobe.com/security/91014620_eusig_wp_ue.pdf (10.7.2013)

Van Diessen, Raymond J. and van der Werf-Davelaar, Titia, *Authenticity in a Digital Environment*, IBM / Koninklijke Bibliotheek Long-Term Preservation Study Report Series, No. 2, IBM Netherlands, Amsterdam, December 2002, http://www.kb.nl/sites/default/files/docs/2-authenticity.pdf (22.2.2013)

Wallace, C., Pordesch, U. and Brandner R., Long-Term Archive Service Requirements, IETF Trust, 2007, pp. 1-17, http://tools.ietf.org/pdf/rfc4810.pdf (9.7.2013)

Wikipedia, XML Signature, June 2013, http://en.wikipedia.org/wiki/XML_Signature (10.7.2013)

XML Advanced Electronic Signatures (XAdES), Technical Specification, ETSI TS 101 903 v1.2.2, April 2004, http://uri.etsi.org/01903/v1.2.2/ts_101903v010202p.pdf (10.7.2013)

XML Advanced Electronic Signatures (XAdES), W3C Note 20 February 2003, ETSI, 2003, http://www.w3.org/TR/XAdES/ (10.7.2013)

XML Signature Syntax and Processing (Second Edition), W3C Recommendation, The Internet Society & W3C, 10 June 2008, http://www.w3.org/TR/xmldsig-core/ (10.7.2013)

Does Digital Rights Management Affect the Mobile Application Market?

Tihomir Katulić
Faculty of Law
Trg maršala Tita 14, 10000 Zagreb, Croatia
tkatulic@gmail.com

Kruno Golubić
University Computing Centre, University of Zagreb
Josipa Marohnića 5, 10000 Zagreb, Croatia
kruno.golubic@srce.hr

Summary

Distribution of digital content is a key aspect of electronic commerce. Digital content industry is threatened by rampant digital piracy. Institutional collective management is slow to adapt to rapid technological change in the digital environment. In the case of mobile software applications, smartphone and tablet operating systems with centralized application repository augment the ability of authors and rights holders to control the commercial exploitation of their works. However, digital rights management technologies used to prevent unauthorized reproduction, distribution and use of protected works can also create unwanted market effects limiting authors and users.

Key words: Digital Rights Management, DRM, competition, intellectual property, mobile applications, application marketplace

Introduction

Intellectual property rights, especially exclusive rights concerning economic exploitation of his work, stand against the widely accepted rules and regulations regarding competition and free market economic behaviour.[1] Even after fifty years of development of the common European market the tension between these two well-established legal disciplines is a much debated issue without definitive conclusions and universally accepted policies [Magnani, Montagnani,84].

In the decades past, there have been many legal and economic arguments concerning the character and scope of application of competition rules on the system of intellectual property in the common European market. In this paper, we will consider a current and important phenomenon – the question of the impact

[1] See References 2-5.

of digital rights management (DRM) technology and its legal position on the development of the mobile application market.

What is DRM?

DRM, or Digital Rights Management stands for technical measures, material (hardware) or immaterial (software) products whose purpose is to allow the legitimate user limited access to protected content in digital form. These technical measures (or technical protection measures, as defined by the articles 11 and 12 of the 1996. WIPO Copyright Treaty) are meant to stop an unauthorized user from copying and distributing protected content in digital form.

In theory there are many competing definitions as to what DRM actually consists of. A few follow for the sake of better understanding of this complex issue:

I) "*Digital rights management (DRM) is a type of server software developed to enable secure distribution – and perhaps more importantly, to disable illegal distribution – of paid content over the Web"* [Rump, 3].

While DRM technology can certainly be deployed via centralized server-client architecture, this is by no means its only *modus operandi*. DRM can be a standalone product or encryption software integrated with the protected content.

II) "*DRM covers the description, identification, trading, protecting, monitoring and tracking of all forms of usages over both tangible and intangible assets"*

This definition is perhaps too broad and does not focus on the main issue – that these technologies are rights management technologies – technologies dedicated to manage the use of protected content on behalf of the author, by the user. Furthermore, these technologies are digital in nature, and relate primarily to protected digital content [Rump, 4].

To conclude, DRM is an encompassing term for several different technologies used to enforce pre-defined limitations on how to access and use protected digital content. The technical protection measures we mentioned earlier are sometimes referred to as DRM technology, however some choose to differentiate between the two by defining technical protection measures as the technology used to control and restrict access as opposed to DRM as technology that relies on technology protection measures to implement these controls and restrictions.[2]

While this technical distinction may be important in identifying actual protection technology, speaking from a legal discourse DRM has historically been regulated as a technical measures and rights management information by the WIPO Internet treaties (articles 11 and 12 of the WIPO Copyright Treaty and articles 18 and 19 of the WIPO Performances and Phonographs Treaties) and the European InfoSoc Directive (article 6 of the Directive 2001/29/EC of the

[2] "Digital Rights Management and TPM", Office of the Privacy Commissioner of Canada, 2006., available at: http://www.priv.gc.ca/resource/fs-fi/02_05_d_32_e.asp

European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society).
Defined in this manner, technical protection measures would consist of using various methods of product activation and access control, such as Internet product activation, using a registration key or a hardware device (*dongle*), as well as using various encryption methods to encrypt data on physical data storage devices and mediums as well as using digital watermarks to identify intended recipient or the document source.
Before we analyze the current framework and assess its impact on the development of the mobile application market, a couple of clarifications and assumptions need to be made.

DRM, intellectual property and competition

Current legal doctrine, to the best of our knowledge, broadly accepts the notion of influencing the application of intellectual property rights through competition principles. Increasingly, intellectual property rights are being curtailed by the application of competition principles [Korah, 432].
This notion is especially present in the *common law* legal doctrine, especially in the United Kingdom, United States [Korah, 433] and other *common law* legal systems, such as those of the former and current *Commonwealth* members. However, this idea is contested in the European continental legal systems, once again illustrating the divide between these two legal disciplines [Korah, 434].
DRM technology has, through broad international and national legislative efforts, as mentioned earlier, become a legally recognized and regulated means of managing the use and distribution of protected content in the digital domain.
Deploying mobile applications, in essence computer software developed by third parties for mobile phones, tablets and other portable computing devices, follows rather different rules then obtaining and licensing software for traditional desktops and laptops under well-established operating systems such as Microsoft Windows, Apple OS X or Linux. Mobile applications and mobile operating systems, most notably Apple iOS, Google Android, Windows Phone and quite a few other legacy (Symbian, BlackBerry) and upcoming operating systems (Mozilla Firefox OS, Ubuntu Mobile etc.) employ a fundamentally different model of application distribution then their desktop counterparts.
Where desktop users usually choose the method of application distribution, and intrinsically the application origin, the mobile users usually have only the option to download and install applications through the official market (such as Windows Market, Google Play or Apple App store).[3] These markets employ a DRM (Digital Rights Management) solution to regulate the way users can ac-

[3] Some mobile operating systems, such as Android or legacy Symbian, allow users to install applications originating from third party marketplaces or other sources.

cess, install and use protected software, make in-application payments to access additional functions and receive periodical upgrades and offers.
Using exclusively official market content (the only available option for Windows Mobile or Apple iOS users) usually means a higher level of quality control and application security, having only one official channel of distribution also means conformity with the less beneficial traits of the DRM protected application delivery system. Furthermore, the fact that an application has been approved to enter the market does not mean that it will remain accessible in the foreseeable future. Once approved, an application can be removed from market at any given time by the market service provider.[4] Revocation of already published applications and preventing certain applications or even types of applications to enter market is an issue of concern to parties other than developers and market service providers.[5]
Finally, the ability of market service providers to revoke a potentially harmful application is instrumental in securing distributed applications.[6] The growing mobile devices market has been experiencing a surge in the development of increasingly sophisticated malware applications on an almost a daily basis.[7]

The role of the DRM technology in the digital content market

Fundamentally, DRM technology serves as an access control tool to protected content. It manages the right to control access to protected content according to the rights of the user. It prevents the unauthorized user from accessing protected content and prevents illegal reproduction and distribution of the protected content. Historically, DRM technology has appeared both as a standalone technical protection measure as well as a part of a broader hardware and software platform (such as the Apple iTunes system).
Concerning the management of protected digital content in contemporary transnational information economy, content enterprises usually adopt one of the two

[4] Mobile application acceptance process for Apple's App Store is based on a profile certificate check approach, which is a standard practice that ensures the integrity of software applications. Reasons for revoking an application from the market can vary from violation of a license agreement to different security concerns due to lost encryption keys. Ability of market owners to revoke application is very important to ensure security of distributed applications.

[5] Removing a successful application from the App Store has attracted the attention of French government on at least one occasion. http://www.reuters.com/article/2013/04/11/apple-france-appstore-idUSL5N0CY42J20130411

[6] Even the purportedly secure iOS ecosystem has been found to harbour at least one malware application that slipped past the curators. http://www.wired.com/gadgetlab/2012/07/first-ios-malware-found/

[7] There are numerous reports regarding the mobile malware development in the security industry. One of the recent research papers analyzes the case of iOS malware corrupting an already inspected application via remote modification of code. See reference 12.

strategies. The content company either develops its own digital distribution channel where it licenses access to its content (like Amazon or Sony), or it adopts an existing distribution channel developed by another company out of specific content business (i.e. the case of music and film studios cooperating with Apple or Google).

It is important to note, however, although it does not in itself present the main topic of this paper, that DRM technology in and of itself presents a specific and important market. Not only do access control systems limit users in accessing and using the protected content in terms of interoperability, preventing the use of content licensed by one company through hardware and software of another, the technology itself presents a product [Magnani, Montagnani,84] and a market where rights holders have an intrinsic motive to adopt the most effective, affordable and secure technological protection measures. Preventing a competitor in obtaining the most effective DRM technology can also present a present a violation of free competition.

Since DRM is legally protected, and it's circumvention is prohibited by a widely accepted legal framework, it is an excellent example to study the effects of DRM technology on application market, and indirectly to assess the impact of legal protection of DRM on competition in this specific market.

Fairplay/iTunes as an example of DRM protected distribution system

As a content distribution system, iTunes allows users licensing and access to music (albums and individual songs), films, television series episodes, radio broadcasts, podcasts etc. Since digital distribution eliminates the need for material data storage disks, effectively provides free shipping and abolishes other costs associated with material products, the licensing fees for digital content are usually lower. Individual songs or broadcasts can often be licensed for less than one dollar or euro, and allow users to license only content they're interested in (not the whole album, season of television series etc.).

This licensing flexibility, combined with lower cost, has arguably been the main reason behind the massive success of iTunes allowing it to become the most popular digital distribution system in history.[8]

Massive success of iTunes has also attracted the attention of national competition authorities.[9] Some have concluded that Apple's use of DRM in creating a

[8] Since April 2004., iTunes has grown from initial 200,000 digital content files to over 20 million individual digital content files including music, podcasts, films, music videos, audio books and over half a million computer programs for Apple OS X and iOS computer and mobile device operating systems. In the same time period the number of individually registered users has reached almost 600 million users, almost 25% of all Internet users in 2013., according to Internetworldstats.com and CNN, available at: http://tech.fortune.cnn.com/2013/06/15/apple-algebra-itunes-asymco/

successful online distribution scheme did not represent a market violation, others disagree. While we do not agree with the opinion of the French competition authority, its analysis of the Apple's practices yields interesting conclusions important for the development of the mobile application market.

As we have mentioned earlier, digital distribution of content eliminates the need for material data carriers, packaging and physical distribution reducing the costs of distribution to a modest fraction of the original cost. Rights holders have eagerly accepted iTunes as a new distribution channel, especially in the face of rampant digital piracy and declining sales of CDs, DVDs and other material media. While Apple has since all but abandoned its proprietary "Fair Play" DRM system, its introduction in 2003 has secured early support of the many of the world's largest rights holder companies and media corporations.

The terms and conditions in the iTunes user's licence agreement have set the standards for the most DRM protected services. Among the most dangerous, from the aspect of consumer protection, is the provision allowing the service provider to limit or deny access to licenced content in the event of closing the iTunes service [Roth, 524]. Furthermore, the user is limited in copying licensed content and authorizing additional devices to access licensed content [Roth, 524]. These terms are still valid for the remainder of the iTunes catalogue, and may relate to licensing of mobile applications from the App Store as well.

How does Apple's behaviour, and indeed the whole iTunes/AppStore/iOS ecosystem affect competition? In order to answer that question the first condition is the need to recognize the relevant market.

Since iOS is Apple's proprietary operating system exclusive to Apple's devices, it is obvious that accepting the use of iOS basically means using Apple's devices as well. If that is the case, why should using the AppStore as an exclusive source for purchase of additional software present a market violation? Is not Apple well within its right, as a competitor in an immensely competitive market (mobile reproduction devices, smartphones and tablets), to choose a business model that binds the hardware (devices) and software (iTunes, iOS, App Store) into a whole eco-system as a best chance of market success? The users can choose from a variety of competitors offering similar devices and services, and, as the French authority asserted in *Virgin Mega* case, Apple's market share concerning multimedia playback devices (iPod) cannot even be considered a dominant position?

The answer is the market creating effect of DRM. Not only does DRM prevent unauthorized access, reproduction and distribution, it also has a market creating effect by preventing users from accessing and using licensed content in a manner of their own choosing. Never before has licensing of content, music, movies

[9] The above mentioned *Virgin Mega v. Apple* case and the opinion of the Norvegian consumer rights ombudsman available at: http://forbrukerportalen.no/filearchive/Complaint%20against%20iTunes%20Music%20Store.pdf

or computer software implied a limitation on the actual type of device being used to access content. By limiting the users to their own hardware ecosystem, Apple has, effectively, created a new market. This is obvious in the case of mobile applications where choosing Apple devices and their iOS system means accepting AppStore as the only legitimate source of new applications. It is less obvious in the case of licensing music or video content, nevertheless Apple's implementation of DRM and its restrictions preventing users from using licensed content on other devices meant a fragmentation of the existing market and the creation of a new niche market – the market of content for Apple devices.

This development in practice meant that unlike most of the competition in the traditional digital content market Apple managed to control the devices and download services monopolizing the market of Apple devices in terms of content distribution. Effectively, Apple has succeeded in using intellectual property protection (copyright on software, as well as patents on hardware and software) as a means to monopolize access to its devices. While it is understandable that competition regulators cannot foresee the future market development, the disparity between the decisions of the European regulators in the cases of Microsoft and Apple illustrates how information technologies can have unexpected economic legal consequences. In hindsight, Microsoft's quasi-monopoly in the desktop and server operating system market has shown to be much less dangerous and competition, coming both from Apple and Google as well as from the Free Software/Open Source community has successfully developed competing products and business models. On the other hand, the company whose devices were just one of the many present in the mobile media player market which in itself at the time of the consideration was a minor market in the broader media appliance market turned out to represent a much more serious challenge to maintaining competition in several emerging markets (mobile phones and applications, digital publishing and media).[10]

To conclude, it can be argued that Apple's behaviour and usage of DRM technology resembles the effects of a tie-in agreement – from the competition perspective, a well-understood practice of selling a product or a service as a mandatory addition to a purchase of another product or service.

The competition practice in the EU (Microsoft Corp v. Commission of the European Communities, 2004) as well as the United States (Jefferson Parish Hospital vs. Hyde, 466 US S2 (1984) has identified five conditions needed to judge a tie-in agreement a market violation. The five-step tie-in test [Schmidt, Hedwig, 183] may or may not still be viable in general terms and may be substituted in the future [Schmidt, Hedwig, 183], but until such time the current criteria consists of: existence of two, from the consumer's perspective unrelated prod-

[10] In 2012. Apple became the company with the highest market capitalisation in the US, worth over 600 billion USD.

ucts, the dominant position of the market player forcing the tied purchase, preventing consumers from purchasing products separately, endangering the competition and absence of legitimate, objective reasons for a tie-in agreement.
From user perspective, Apple's ecosystem and the restrictions regarding the types and functionalities of applications allowed on its smartphones has from the start been received with attempts to circumvent or disable technical protection. This procedure, colloquially known as "*jail-breaking*" has been itself an object of official inquiry and judicial procedure, at least in the United States.[11]

The mobile bazaar

In the seminal 1996 essay "The Cathedral and the Bazaar", a decade before the appearance of earliest modern online mobile application stores, the veteran open source evangelist E. S. Raymond described the rise of two distinct models of application distribution, one governed by a single hierarchical entity (the cathedral) and one open to parallel input from different, competing or cooperating sources (the bazaar) [Raymond, 3].
Where Apple's model is obviously one of a cathedral, where the market service provider ultimately decides to accept or reject publishing software and where users have no legal recourse to challenge the market service provider's decisions, Google's Google Play model is, in line with open source heritage the company draws from, a practical exercise of the bazaar model.
Allowing users to choose between a curated Google Play market yet retaining the option to allow installation of applications originating from other, independent markets or directly from the World Wide Web or a data storage device, Google has adopted a model that fosters competition and openness. Even if an application does not qualify for distribution through Google Play market (be it for quality control reasons, security etc.) the user still has an option to install it through another independent market, directly from the manufacturer's website or by locating and installing it manually.
The phenomenal success of the Android mobile operating system, which accounts for more than three quarters of all smart mobile phones and more than a half of all tablet computers owes in no small part to the bazaar model adopted

[11] In 2010., the US Court of Appeals for the 5th Circuit, No.08-10521 MGE UPS Systems INC v. GE Consumer and Industrial Inc etc. decided that "*The owner’s technological measure must protect the copyrighted material against an infringement of a right that the Copyright Act protects, not from mere use or viewing*", effectively establishing *jail-breaking*, at least in the case of the iPhone, a fair use. In addition, in a previous statement by the federal Copyright Office with regard to Apple's claim of copyright protection over the encryption software in the iPhone's bootloader the Copyright Office concluded: "*While a copyright owner might try to restrict the programs that can be run on a particular operating system, copyright law is not the vehicle for imposition of such restrictions.*" In other words, copyright and even legal protection of technical protection measures cannot be used to limit the users with regard to the way they use their devices.

by Google, and by example followed by the major manufacturers of Android mobile phones.[12]

Conclusion

Considering the success of mobile application eco-systems developed around the App Store/Marketplace model and the central role of DRM technology in all but one of the competing systems it is safe to conclude that DRM technology can exhibit a strong anti-competition effect.

It is obvious that open systems allow more user freedom and foster competition. While absence of a closed curating system may present security risks and overall lower levels of quality control, market forces ensure that users can choose quality and advanced technical capabilities unavailable in the closed markets due to technical limitations imposed by the proprietary systems or prohibitive policies of the market service provider enforced by its DRM.

However, this is not the case against adopting and further refining DRM technology. The technology itself is neutral and can be put to constructive use as well. Effective DRM technology can enable individual authors – artists, journalists, programmers – to develop and monetize their work online without the need for intermediaries like publishing houses or collective rights management societies. Even the closed off application markets of today with their anti-market DRM implementation are often a more affordable and effective choice for today's authors. DRM and developing content management systems hold the promise to finally turn the tide of rights management from collective to individual – for the first time since the invention of the printing press.

References

1. Becker, Eberhard; [et al.]. Digital Rights Management: Technological, Economic, Legal and Political Aspects. Berlin; New York : Springer, 2003.
2. Czapracka, Katarzyna: "Intellectual Property and Limits of Antitrust: A Comparative Study of US and EU Approaches", Edward Elgar Publishing, 2010.
3. Drexl, Josef: "Research Handbook on Intellectual Property and Competition Law", Edward Elgar Publishing, 2010.
4. Ehlermann, Claus-Dieter, Atanasiu Isabela: "European Competition Law Annual 2005: The Interaction between Competition Law and Intellectual Property Law", Hart Publishing, 2007, etc.
5. Ghidini, Gustavo: "Intellectual Property and Competition Law: The Innovation Nexus", Edward Elgar Publishing, 2006.
6. Korah, Valentine. The Interfrace Between Intellectual Property Rights and Competition in Developed Countries.// *SCRIPTed*. 2 (2005), 4; 430-443.
7. Magnani, Paola; Montagnani Maria Lilla. Digital Rights Management Systems and Competition – What Developments Within the Much Debated Interface Between Intellectual Property and Competition Law?, // *IIC* 39 (2008): 83-105.

[12] Exact number is 79.3% of the global market share in the second quarter of 2013. Source: IDC Worldwide Mobile Phone Tracker, August 7, 2013

8. Raymond, Eric Steven: "The Cathedral and the Bazaar", p. 3, available at: http://www.catb.org/esr/writings/homesteading/cathedral-bazaar.
9. Roth, Monika. Entering the DRM-Free Zone: An Intellectual Property and Antitrust Analysis of the Online Music Industry. // *Fordham Intellectual Property Media and Enterprise Law Journal*. 18 (2007): 515-540.
10. Rump, Niels: "Digital Rights Management: Technological Aspects", p. 3, "Digital Rights Management: Technological, Economic, Legal and Political Aspects", edited by Becker, E., LNCS 2770, Springer, 2003.
11. Schmidt, Hedwig. Competition Law, Innovation and Antitrust: An Analysis of Tying and Technological Integration. Cheltenham : Edward Elgar, 2009.
12. Wang, Lu, Chung, Lee: "Jekyll on iOS: When Benign Apps Become Evil", Proceedings of the 22nd USENIX Security Symposium, August 2013.

Mutation of Capital in the Information Age: Insights from the Music Industry

Sergej Lugović
Polytechnic of Zagreb
Vrbik 8, Zagreb, Croatia
lugovicsergej@gmail.com

Sonja Špiranec
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
sspiran@ffzg.hr

Summary

The question this paper addresses is whether more variety in meaning of capital could increase the value of capital as one of the factors of production. Methodology used is based on observation, and in this stage of research of this phenomenon we believe that detailed observations are needed with aim to provide enough qualitative data for further development of qualitative methods. Based on a case study from the music industry, the authors have tried to determine whether the extensive use of technology, relationship, knowledge and emotions can be used as a resource for value production in today's economy? These four factors, under complex system analyses of their relationship, could provide enough variety to respond to increasing variety (diversity) developed in music industry environment. In the first part of the paper the on-going rise of importance of intangible values and their relationship with tangible one is discussed. After that, value theories are observed and their relationship to information science is being examined. In the last section, the authors will assert that increasing complexity, coupled (combined/conjoined) with faster changes in the music industry should directly be related to the development of new factors of production used to generate value which can be exchanged on the market. The proposed changes will be tested on a case study.

Key words: general system theory, capital mutation, information economy, music industry, law of requisite variety, value theories

Whereas at one time the decisive factor of
production was the land, and later capital –
today the decisive factor is increasingly man
himself, that is, his knowledge.
Pope John Paul II (1991). Centesimus Annus

Introduction

Recent theoretical considerations in the field of economy have elicited two, at first sight opposite thoughts. One is that managers are only responsible to deliver profits, while the other one focuses more on values, which are intangible and hard to measure. These two different thoughts come from two of the probably most important economists in the last century: John Maynard Keynes and Milton Friedman.

In his book "The General Theory of Employment, Interest, and Money", Keynes says:

> "Thus if the animal spirits are dimmed and the spontaneous optimism falters, leaving us to depend on nothing but a mathematical expectation, enterprise will fade and die;—though fears of loss may have a basis no more reasonable than hopes of profit had before"[1]

Such thoughts introduce a whole new field of research in social sciences, where everything cannot be mathematically measured. To ask what is an animal spirit, will probably be the same as asking how much does a hard drive with 5 years of family photos cost and how its owner would estimate its value. It has a value which probably cannot be expressed in mathematical terms, but only in relationship with something else. In his probably most cited sentence, "The Social Responsibility of Business is to Increase its Profits"[2]. Milton Friedman suggests that only responsibility of business activity is aimed toward making profit.

As profit is tangible, with such a view management of business is only reduced to tangible aspects.

One could say that if business increases its profit, it has a good animal spirit in business organization. Therefore, intangible and intangible aspects of business operations are not opposite, they are complementary phenomenon. To put them in such constellation could give us much broader and better framework for analyzing, researching and measuring the success of companies. Hence, we have to analyze things which are tangible, and their relations, which are executed by humans and/or machines. As relations are intangible and things are tangible,

[1] Keynes, John M., *The General Theory of Employment, Interest and Money*, London, Macmillan, 1936, pp. 161-162.

[2] Friedman, Milton, The Social Responsibility of Business is to Increase its Profits, *The New York Times Magazine,* New York, 13 September 1970., http://www.colorado.edu/studentgroups/libertarians/issues/friedman-soc-resp-business.html

and humans are subjective – we could look on them as a complex system with dynamical relationships.

Currently the only existing measurement of the company value is what somebody is willing to pay for the company. This process of evaluation is conducted on the free market. Buying shares of one of the Standard and Poors 500 (S&P 500) companies, only 20% is measurable in tangible values, and another 80% is intangible.[3] Such a transaction is based on intersubjectivity of a buyer and a seller.

In order to provide a better picture, we made a brief analysis of the Forbes top ten most innovative companies today whose information can be found on the NASDAQ.

Table 1.

Company	Market Cap (20.09.2013)	Total Assets	Assets Report Date	Ratio Assets / Market Cap	Average Assets / Market Cap	Actuall Difference in mlrd. USD
	In mlrd.	In mlrd.			26,63%	**767,406**
Salesforce.com	31,814	5,528	1.31.2013	17,38%		26,285
Amazon.com	142,575	32,555	12.31.2012	22,83%		110,020
Intuitive Surgical	14,691	4,059	12.31.2012	27,63%		10,632
Apple	429,083	176,064	9.29.2012	41,03%		253,019
Google	299,182	93,798	12.31.2012	31,35%		205,384
Monsanto	56,973	20,224	8.31.2012	35,50%		36,749
Celgene	61,082	11,734	12.31.2012	19,21%		49,348
InfoSys	27,718	8,539	3.31.2013	30,81%		19,179
FMC Technologies	13,539	5,902	12.31.2012	43,60%		7,636
Starbucks	57,369	8,219	9.30.2012	14,33%		49,150

Nidec, which is on the list, is excluded because it's a holding company.

Sources: http://www.nasdaq.com, data retrieved on 20.09.2013; http://www.forbes.com/special-features/innovative-companies.html, data retrieved on 20.09.2013;

This table indicates that in only ten companies the actual difference between their assets value and their market value is around 767 billion Euros.

Shapiro and Varian[4] defined term *information delay* as an important factor in the network economy. What they say is that information could increase or re-

[3] Ocean Tomo's Annual Study of Intangible Asset Market Value – 2010, *Ocean Tomo Intellectual Capital Merchant Banc* press release, Chicago; 4. 2011, http://www.oceantomo.com/products andservices/investments/intangible-market-value

[4] Shapiro, Carl, Varian, Hal R., Information Rules, Harvard Business School Press, Boston, Massachusetts; 1999, p. 4 (Chapter 1, Information Economy)

duce its value according to amount of delay in time. Smaller delay – higher value, bigger delay – smaller value.
Using information of market value, as a measurement and evaluation tools of company activities, we are addressing one problem – delay of information, and value of this information, as a delay from the moment when some activity is done and that activity reflected in share price could be quite long.
The Interplay of value, relationship, humans, machines and profit, could be perceived as a starting point to produce something what is tangible and could be exchanged on the market, by the price which is nothing else but the information about the value that is exchanged. So, to produce something, along raw materials, we have to use our thoughts as well, which are intangible. Therefore anything that enters economy system of exchange already has both, tangible and intangible coded into it. The way how we manage such a complex system of dynamic relations could have direct impact on market value of business organization.
The aim of the paper is to present potential interplays between information and economy. and specifically to determine whether the extensive use of technology, relationship, knowledge and emotions can be used as a resource for value production in today‘s economy.

Theoretical background

"System analysis of a business enterprise encompasses men, machines, buildings, inflow of raw materials, outflow of finished products, monetary values, good will and other imponderables it may give definite answer and practical advice." – Ludwig Von Bertalanffy[5].
Not to leave Bertalanffy alone in system approach, we could cite Ronald Coase[6], who is calling economists to look the economy as a whole system as an object of study, how one part is related to another, how they are interrelated and how they actually work together. Further on, he states that we have to develop sensible analysis of comparison of the additional production and rearrangement of activities with transaction costs.
Rearrangement of activities leads to new relationships which could give new meanings, and meanings define new properties which are then analysed and applied in complex system of dynamic relationship that could lead to new factors of production.
So it looks that they are talking about same thing. About "Whole". About a System. About a dynamic and complex relationship in that system.

[5] Bertalanffy, Ludwig Von, General System Theory: Foundations, Development, Applications George Braziller Inc, New York; January 1984, 9th printing p. 196

[6] Coase, Ronald, Why Economics Will Change, *Remarks at the University of Missouri,* USA Columbia Missouri; April, 2002, http://www.coase.org/coaseremarks2002.htm

In his book Cybernetics, Norbert Wiener[7] proposed the lesson that any organism is held together by possession of means for acquisition, use, retention and transmission of information and that the means of communication are the most important and most effective. As the book was written more than 50 years ago, it is interesting to see what today scientist think about such a role of information in today's economy. Wolfgang Hofkirchner in his recent paper[8] is highlighting the importance of having a Unified Information Theory. Here he presents examples of concrete-universal concept as transformation driver in development of capitalism, which are *"sometimes regarded as mutations into different economic system"*. In the same example he states that such developments did not replace capitalist principles fully, but modified the essence of it, referencing to the latest notions such as informational capitalism. So, from the proposal of Professor Wiener in 1949 about importance of information for organization, which also may refer to business organizations, to the mutation of capital principle – we could conclude that such a mutation is directly related to the development of the information science.

In what direction are such mutations appearing? For example, according to Hamel[9], instead of control we should build up a trust. It is much cheaper to build a trust, then a system for controlling employee's activities. At the same time, building a trust is not a simple formula we could pull out from some classical economic books. It is heavily dependent on intangible part of complex system and involved/associated relationships discussed in this paper. Could you imagine increase of profitability only by smiling and looking into the eyes of employees, with nothing to hide, driven by higher moral values and positive emotions? It would not cost you anything but will increase profitability.

Along with not investing money on the control equipment the quality of working atmosphere is being increased which then leads to better working results and lower costs. So company assets are decreased and liability is increased but we still have better results. It is completely opposite to classical capital formula where capital is assets minus liabilities.

There are a lot of such activities in today's economy, and they certainly have the role in productivity, but somehow they are left outside of the neoclassical economist thoughts.

---

[7] Wiener, Norbert, Cybernetics: Or Control and Communication in the Animal and the Machine, *MIT Press,* Cambridge Massachusetts; 1948, p. 187

[8] Hofkirchner, Wolfgang, Emergent Information Some System – Theoretical Considerations About An Integrative Information Concept, *International Journal "Information Theories and Applications",* Vol. 18, Number 1, Institute of Information Theories and Applications FOI ITHEA, Sofia Bulgaria; 2011

[9] Hamel, Gary, Moon Shots for Management, *Harvard Business Review*, Boston Massachusetts; February 2009

Those are things which are intangible, but they are not labour, nor land, so only what they could be is capital, according to classical factors of production[10] .A kind of a new mutated capital, which heavily depends on information instead on tangible matter. Such a transition of meaning of capital resulted in change of properties of the capital. This transition happened through complex process of interaction with other entities in system, in different situations, over the time. What Bernard Lietaer [11]says – for clapping hands we need two hands, one is not enough, and maybe these two hands could be tangible and intangible, going together regardless their possibility to quantify or not.

To develop scope, or lens on how to look at intangible, which is having a role in the process of production of goods or services traded on the market, it is important to understand theories of how value is created. Three important theories can be used as a starting point. One is the theory of *value chain* by Michael Porter[12] who explains how company produce value by executing the steps of activities in order, where total added value is more than sum of the independent steps or activities. Again, we could refer to the Global System theory which addresses the concept of "Whole", which has more value than sum of its part. To adopt this theory in information age John Sviokla and Jeffry Rayport[13] developed a concept called Virtual Value Chain, in which, along Porters value chain, an organization should add activities which should produce value out of using information. Again, the sum of total value is higher than the sum of independent steps. For those activities object is information, or bits, or something intangible that is gathered, organized, selected, synthesized and distributed. This theory aligns with proposed view that tangible and intangible should be perceived together. A third theory about value is the *Stake Holders Theory*, written by R. Edward Freeman[14], which puts in the focus and other stakeholders in company system environment to produce value. Along resources based view and market based view this theory adds socio- political view. To produce maximum value, companies or their directors have to include all three views in organizing their operations and steering organization system. This theory focuses on relationship inside the system where one company operates by expanding the view outside of legal definition.

[10] Samuelson, Paul A.; Nordhaus, William D., Economics, Tata McGraw Hill, 18th edition, New York; January 2006

[11] Ulanowicz, Robert E.; Goerner, Sally J.; Lietaer, Bernard; Gomez, Rocio, Quantifying sustainability: Resilience, efficiency and the return of information theory, *Ecological Complexity,* 6, no. 1, Elsevier, Philadelphia, 11.29.2008

[12] Porter, Michael E., Competitive Advantage, *The Free Press,* 6.1.1998, p 36

[13] Rayport, Jeffrey F.; Sviokla, John J., Exploiting the Virtual Value Chain, *Harvard Business Review,* Boston Massachusetts; November 1995

[14] Freeman, R. Edward, Strategic Management: A stakeholder approach, Pitman, Boston; 1984

All three theories have addressed intangible part of doing business, and suggest how to produce more using tangibles. So, combining them may be perceived as a prerequisite for more successful business operations.

Table 2. System view of theories

Theory	System view
Value Chain	Whole
Virtual Value Chain	Tangible and Intangible
Stakeholders theory	Dynamic Relationship

Today's modern economies create their revenue or income based on services, (it was 68% in 2005, and industry was only 30%.[15] Why is this important? Because, services are tangible only if sold, and their basic ingredient is information. So actually, dependent on how information is processed, will result in profit and market value.

If we look at the relationship of tangible and intangible as a system, and then use this system as a mean of production then, whatever we produce and enter the economic system has intangible and tangible in it.

So it is important to look them as one, not as separate entities, and even more important, we have to focus on how it is produced and what are the relations in that process.

In the next chapter, the described phenomenon and new perspectives in analyzing intangible dimensions of value creation will be illustrated in the domain of music industry.

Contemporary definitions of capital and the Music industry

The Music industry was one of the first which moved from industrial age to digital or information age[16] so it is attractive playground for empirical research.

The theoretical frameworks discussed in the previous chapter suggest that it is not only necessary to concentrate on results, but that more analysis is needed regarding how something is done. Therefore, several questions need to be addressed:

- How we use technology available to produce value?
- How organization is driven by emotions?
- How we use relationship to produce value?
- How we apply knowledge in this process?

Let us give a brief example.

If we compare two companies, A and B, both owning a catalogue of 10,000 songs in similar genre and with similar popularity, and company A is not inter-

[15] Memedovic, Olga; Lapadre, Lelio, Structural Change in the World Economy: Main Features and Trends, United Nations Industrial Development Organization, Vienna; 2010

[16] IFPI – Digital Music Report 2012, London, 2012, http://www.ifpi.org

ested in using technology to promote and sell them digitally, corporate culture is not driven by positive experience and emotional attachment to the music they have in catalogue, they are not interested in exploring value of relationship they have built over the years, and they are not acquiring new knowledge how to conduct the business in digital environment, while company B is doing opposite, using technology, driven by positive experience and emotions, using relationship to the maximum and developing knowledge. It is not hard to answer to question which one will have higher market value. This is just a theoretical example, and it is used as an explanation, not as a scientific proof. Therefore, it is reasonable to address the basic classical economic view on factors of production, i.e. and they are Land, Labour and Capital[10]

In his dictionary, The Economist[17] is calling capital money or assets put in economic use, and it is described as one of the four ingredients of economic activity, included in factors of production, such as land, labour, capital and enterprise, where enterprise is defined as animal spirit of entrepreneur who is able to mix factors of production to produce something valuable. In other sources capital is described as a stock of resources employed in the production of goods and services and that it has many meanings[18]. This variety of approaches and viewpoints in defining the concept of capital leads to the conclusion that more research is needed about the actual definition of the term, especially in the light of the fact that essence of capitalism mutated under the influence of information as means of production.

The classical theoretical framework, where factors of production are determined as land, labour and capital may be questioned based on the following analogy: if land is space, labour is time and capital is energy, then we may assert that labour as a factor of production is covering relationships between employer and employee, but creating the playground for such a relationship then could be called capital.

For the purpose of this paper labour as factor of production has to be interpreted as an individual skill set, personal values systems, personality and time available for organization or corporations to use to produce values. Capital is everything else. If we take two organizations, whose employees interact with the aim to produce some products or services, if some of the employees left the company where they are employed, they could take their time with them, but two companies still have their relationship as capital as one of the factors of production.

To put it differently, everything what is not space (or land) where labour spend their hours (time), could be called capital. And capital could be called energy

[17] Economics A-Z terms beginning with C, *The Economist,* London, retrieved 20. 08. 2013, http://www.economist.com/economics-a-to-z/c

[18] Capital and Interest, *Encyclopedia Britanica (online edition),* retrieved 20. 08. 2013, http://www.britannica.com/EBchecked/topic/93850/capital-and-interest

which is needed to produce value exchangeable on the free market. So everything what is used along labour and land, to produce product or service, could be called capital. Following idea of researching process instead of its outputs, leads us to directions suggested by Ronald Coase[19]:

"The fact of the matter is that economists commonly obtain their theories in the study of industrial organization (and probably elsewhere) as a result, not of examining what actually happened but by thinking about it."

In other words, profit happens, results happen, but why and how it happens is more important and should be researched more in depth. In the case of conceptualizing the music industry system, we suggest to adhere to the classical economic formula of factors of production, focusing on capital. The question is not what the results of using capital to produce value are, but how it is used in complex system, what impacts have positive emotions in the organization, and how to use technology to achieve better results. Along technology and emotions, relationship and knowledge should be researched as factors of production in today's information dependent economy.

So, how then do we have to look at music industry?

Explaining how classical music industry works[20] it is defined that the Music industrial system involves aesthetic and industrial production embedded in economic production functions and intermediaries that support different phases. They also stated that production of music is more number of intersecting networks where ideas circulate among firms and individuals then linear activities. Furthermore, the authors stated that the music industry is dominated by only 4 companies, and even in the time of the writing of this paper two of this four merged. So we have Music System which is dominated by only 3 companies. To address such a position in the system we would like again to cite Ludwig Von Bertalanffy [21]and his book General System Theory; "if only few or competing pair are left, conflicts become devastating to the point of mutual destruction"

Hence, the system is in a state of risk of devastation and destruction. Still, it can be rejuvenated; by mutated capital, build on information, resulting in "how" view, instead of "what" view to related production factors, i.e. capital.

What we are witnessing is a major development in the music system or environment of the music organization companies in the last 20 years. Good sum-

[19] Coase, Ronald, The Conduct of Economics: The Example of Fisher Body and General Motors Journal of Economics & Management Strategy, 2006, pp. 255-278

[20] Power, D.; Hallencreutz, D., Competitiveness, Local Production Systems and Global Commodity Chains in the Music Industry: Entering the US Market Regional Studies, 2007, pp. 377-389

[21] Bertalanffy, Ludwig Von, General System Theory: Foundations, Development, Applications, George Braziller Inc. New York, January 1984, 9th printing, p. 48

mary of those developments are underlined by Knowles[22], who described these changes as follows:

- democratization of the tools for production,
- democratization and diversification of the means for distribution,
- the rise of social networking sites connecting producers and consumers and listeners with shared tastes,
- the infinite catalogue and 'The Long Tail',
- an exponential increase in the availability of 'free music' through illegal file sharing,
- the arrival of large capacity of portable media devices.

Basically, change in the environment of music system, gave us a lot of variety, and here Ron Ashby's "Law of requisite variety" applies[23] , who claims that in order to control variety , variety is necessary. To answer such a development of the variety in music system, organization has to develop new set of business functions which will respond to such a variety. If the system keeps to be ruled by only 3 companies, regarding the General system theory, it is a matter of the day when it will be destructed. But if the new production factors are developed as a response to the increasing variety in surrounding music system, then music system could manage its existence much better. Their variety should come as an output from analyzing complex system of dynamic relations of tangible and intangible, giving us – depending on how deep we are digging – almost unlimited source for system analysis and healthy infrastructure for value production.

Case study

"The primary way a researcher can investigate an educational organization, institution, or process is through the experience of the individual people, the "others" who make up the organization or carry out the process."[24]

To produce the value, we need the factors of production: labour, land and capital. But, meaning /understanding/subject/denotation of the capital mutated over the time. Could we use technology, relationship, knowledge and emotion as a substitute for financial and tangible capital?

This is phenomena we would like to observe, with purpose of gaining knowledge for further research. At this stage only qualitative observing methodology is used.

[22] Knowles, Julian D., A Survey of Web 2.0 Music Trends and Some Implications for Tertiary Music Communities, *In Proceedings National Council of Tertiary Music Schools Conference 2007.* Music in Australian Tertiary Institutions: Issues for the 21st Century, Queensland Conservatorium Griffith University, Brisbane, Australia, 2007

[23] Ashby, W. Ross, An Introduction To Cybernetics, Chapman & Hall Ltd, London, 1957, p. 206

[24] Irving Seidman; Interviewing as Qualitative Research; A Guide for Researchers in Education and the Social Sciences, Third Edition Teachers College, Columbia University New York and London

To find it out, we interviewed Eli Goldstein from Soul Clap (Boston, USA), musician, DJ and label owner. Soul Clap as a project in the last 7 years has been building quite extensive business operations, including music production, record company and DJ-ing all around the globe. In 2012 they had approximately 120 gigs. They have a major act on leading global festivals.

We have to clearly underline limitations of such an observation because it's based only on one interview. This is the subject of research of one of the authors PhD thesis and more interviews will be done. This observation could be looked as an experiment or preliminary pilot study, which could be used as a base for further development of more robust scientific apparatus.

As an introduction, the theory discussed above was presented to Eli. Usage of the technology, emotions, relationship and knowledge could be looked/considered as the factors of production and extensions of the meaning of capital used to produce the value.

Here is the transcript of the interview[25].

Eli and his partner Charles along with DJ-ing back in 2006 started an online blog called Soul Clap, where they used to write about the music they liked, and in that way promote it to the people who share passion for the similar music taste. Soon after the blog they introduced a regular podcast, which was in the early beginning of the podcast days. (What is also important to state, and it is not from the interview but from the researcher observation, they used Ajax technology as soon as it came out on their website. This positioned them as somebody who understands technology well and knows how to use it. Web site was fast and very user orientated.) As Eli says, they had a good knowledge about the music and their passion for the original house music sound was a little bit different than what was considered to be, at that time, mainstream on the independent electronic music blogs around the globe. So, combining knowledge about music, emotionally driven to share it by using advanced technologies for that time such as AJAX and podcasting, they created website which was attractive to visitors. As visitors started to come back for more of what Soulclap had to offer, slowly fan base started to rise, and relations with fans and similar music organization started to develop. As Eli stated, this relationship built through their web site was a foundation for promotion of their music production. Such a development of the fan base and popularity of the web site encouraged Eli and his partner to leave the business they had started with a third partner, which was also related to the music event promotion and mobile DJ-ing. Guided by positive feedback as high emotional driver they have decided to go to the studio and produce only the music, without being worried about business operations they use to run.

[25] Original audio recording available on request

This situation had a huge impact on their positive emotions which they channelled to produce even better music, by using available technology. When they got enough material they started looking for a company which will release that music. Apparently, as they were active in Boston as a DJs, their friends were starting a record label Air Drop at that time, Eli and his partner decided to release music on their friends' label, since they have had a relationship with them before. At the same time their music production skills have developed and knowledge level rose, so combining deep knowledge about old disco records, they started to do re-edits (updated versions of the old tracks). Through their web site, and building relationship with other music organizations, they have developed a relationship with Wolf and Lamb as well, music organization from Brooklyn, New York. As both of them shared the same emotions for such a musical format and style (disco re-edits) they started to develop collaboration. Eli stated that in that phase Skype had a very big role for the development of this relationship, enabling them to build a relationship with no cost at all.

In 2009 their re-edit was released on Wolf and Lamb. Also followed the podcasts for Wolf and Lamb website. So to make it all happen till here, they have only used technology, relationship, emotions and knowledge. It is also important to say that those four factors in their mutual interaction multiplied produced value. One lead to another. In 2009 Airdrop organized a small tour for Soul Clap in Europe, and this was actually their first exposition for global electronic dance music market. An outcome which really surprised Eli and Charles was that people in European clubs recognized music from the Wolf and Lamb podcasts they have done. They have been asked to play tracks from those podcasts and that kind of music which was not usually played in clubs in Europe, especially with the knowledge and style of DJ-ing they developed in US. Probably the first big appearance was in Berlin club Watergate, which is accepted as an electronic dance music institution globally. What makes this important is that they got booked using their existing relationship with DJ Heidi who was just at that time starting her own night in the same club. As their DJ set was well received cause of deep knowledge of rare and forgotten records, re-edited just for purpose of playing it in their own DJ sets.

They were immediately approached by few bigger booking agencies in Europe. This opened up the next step in their operations and put them in higher league of performers with higher income. When they got back to US they agreed upon doing all business using Skype with no costs. By then, their value system has developed, and it has involved two US companies and booking agency in Europe, and all of them were sharing technology, relationship, emotions and knowledge to create value for everybody involved. Eli stated that as all of these factors now used by bigger system are starting to multiply. In 2011 Soul Clap and Wolf and Lamb together done DJ Kicks compilation (one of the best-selling DJ mix compilations in the world) and again they put together knowledge(about music), relationship (with network of music organizations), emotions (passion

for the music) and technology (to promote with low cost). For the compilations they have selected the artists they were working with, what actually gave them additional value and expanded value system they operate in. After that compilation in 2011 they had about 120 international gigs around the globe.
Interviewer asked – is it important to keep all of these 4 factors in balance, to keep the system growing (if you do not develop one of the factors, could the whole system be affected) Eli answered positively and added that he experienced that a lot of musicians were not able to keep up with growth just because of not developing it in a balanced way. To conclude, in the interview Eli was asked about approximate number of gigs in 2011 and 2012, and did Soul Clap along with using technology, relationship, emotion and knowledge do some major financial investment in their business operations. He completely agreed that building such an extensive business operations was done only by using those four factors without financial investment, and that they could even make more money out of gig as demand is bigger but also price per gig rises, but they are very aware of the risk of misbalancing those factors because of extensive touring. But the income they have generated opened up the new business opportunity which can grow even further.
As a conclusion Eli says: "It is still business, we have to remember it is business, but embracing intangibles along, such as passion, love and fun, and not focusing on business side so much."

Conclusion

At one hand we have situation where complexity of relationship between tangible and intangible in one system, could increase variety of the system, and by this, better respond to development of variety in environment where business is conducted. Also, regarding the conducted case study we open Pandora's Box of what is new mutated capital as a part of factors of production, and how it could be used in the information age to produce value, products and services which are then exchanged on the free market for real money. We hope that phenomena researched in this paper, that money could be made without actually having money as a capital investment, is interesting enough for further research and constructive criticism. It is not that this phenomena is new, it has existed in the information age from the beginning, and as we know Google and Facebook started as a projects in student rooms as well. Writing this paper and conducting research we found a huge variety in the meaning of capital. And this is something what should be researched deeper, from both objective and subjective perspective. The aim was to propose new research framework, which takes into consideration four factors – technology, emotions, knowledge and relationship – which could be analysed as capital in the age of information, and how they work in complex system of today's business. Authors understand that this is preliminary small scale research, and it has limitations because of only one interview subject. Research should incite the development of precise tools of

measuring how such factors are actually used, including exploration, description, explanation and valuation scientific methods.

Acknowledgments

Authors would like to thank Eli Goldstein from Soul Clap for taking part in this experiment.

References

Ashby, W. Ross, An Introduction To Cybernetics, Chapman & Hall Ltd, London, 1957, pp 206

Bertalanffy, Ludwig Von, General System Theory: Foundations, Development, Applications George Braziller Inc, New York; January 1984, 9th printing pp 196

Bertalanffy, Ludwig Von, General System Theory: Foundations, Development, Applications, George Braziller Inc. New York; January 1984, 9th printing, pp. 48

Capital and Interest, *Encyclopaedia Britannica (online edition),* retrieved 20.08.2013, http://www.britannica.com/EBchecked/topic/93850/capital-and-interest

Coase, Ronald, The Conduct of Economics: The Example of Fisher Body and General Motors Journal of Economics & Management Strategy, 2006, pp. 255-278

Coase, Ronald, Why Economics Will Change, Remarks at the University of Missouri, USA Columbia Missouri, April, 2002, http://www.coase.org/coaseremarks2002.htm

Economics A-Z terms beginning with C, The Economist, London, retrieved 20.08.2013, http://www.economist.com/economics-a-to-z/c

Freeman, R. Edward, Strategic Management: A stakeholder approach, Pitman, Boston, 1984

Friedman, Milton, The Social Responsibility of Business is to Increase its Profits, *The New York Times Magazine*, New York; 13. September 1970, http://www.colorado.edu/studentgroups/libertarians/issues/friedman-soc-resp-business.html

Hamel, Gary, Moon Shots for Management, *Harvard Business Review*, Boston Massachusetts, February 2009

Hofkirchner, Wolfgang, Emergent Information Some System-Theoretical Considerations About An Integrative Information Concept, *International Journal "Information Theories and Applications",* Vol. 18, Number 1, Institute of Information Theories and Applications FOI ITHEA, Sofia Bulgaria, 2011

International Federation of the Phonographic Industry – Digital Music Report 2012, London, 2012, http://www.ifpi.org

Keynes, John M., The General Theory of Employment, Interest and Money, London, Macmillan, 1936, pp 161-162.

Knowles, Julian D., A Survey of Web 2.0 Music Trends and Some Implications for Tertiary Music Communities, *In Proceedings National Council of Tertiary Music Schools Conference 2007.* Music in Australian Tertiary Institutions: Issues for the 21st Century, Queensland Conservatorium Griffith University, Brisbane, Australia, 2007

Memedovic, Olga; Lapadre, Lelio, Structural Change in the World Economy: Main Features and Trends, United Nations Industrial Development Organization, Vienna, 2010

Ocean Tomo's Annual Study of Intangible Asset Market Value – 2010, Ocean Tomo Intellectual Capital Merchant Banc press release, Chicago, 4. 2011, http://www.oceantomo.com/productsandservices/investments/intangible-market-value

Porter, Michael E., Competitive Advantage, The Free Press, 6.1.1998, pp. 36

Power, Dominic; Hallencreutz, Daniel, Competitiveness, Local Production Systems and Global Commodity Chains in the Music Industry: Entering the US Market Regional Studies, 2007, 377-389

Rayport, Jeffrey F.; Sviokla, John J., Exploiting the Virtual Value Chain, Harvard Business Review, Boston Massachusetts, November 1995

Samuelson, Paul A.; Nordhaus, William D., Economics, Tata McGraw Hill, 18th edition, New York, 1.1.2006

Shapiro, C.; Varian, H.R., Information Rules, Harvard Business School Press, Boston, Massachusetts, 1999, pp 4 (Chapter 1, Information Economy)

Ulanowicz, Robert E.; Goerner, Sally J.; Lietaer, Bernard; Gomez, Rocio, Quantifying sustainability: Resilience, efficiency and the return of information theory, Ecological Complexity, 6, no. 1, Elsevier, Philadelphia, 11.29.2008

Wiener, Norbert, Cybernetics: Or Control and Communication in the Animal and the Machine, *MIT Press,* Cambridge Massachusetts, 1948, pp. 187

LANGUAGE TECHNOLOGIES

Use of Corpus Analysis Tools in Medical Corpus Processing

Evelina Miscin
College of Business and Management 'B.A. Krcelic'
V. Novaka 23, Zaprešić, Croatia
evelinamiscin@yahoo.co.uk

Summary

The aim of this paper is to show the use of three corpus analysis tools – TermeX, Simple Concordance and Collocation Extract - in processing of medical corpus to obtain collocations. Each tool had its function and will be dealt with separately. TermeX tool was used for obtaining a list of most frequent nouns and the analysis of their frequency. Simple Concordance tool was used to get concordances and for manual extraction of collocations. Collocation Extract tool was used for extracting remaining collocations with a greater distance between a collocate and a node and for determining Log Likelihood and Mutual Information of these collocations. All the data were helpful in determining the most frequent collocations necessary to improve collocational competence of users of medical English.

Key words: corpus analysis tools, collocations, corpus linguistics, medical corpus

Introduction

This paper deals with the use of corpus analysis tools in medical corpus processing in order to obtain collocations. Three tools were used: TermeX, Simple Concordance and Collocation Extract. Each tool will be described separately together with its role in extraction and processing of collocations.

Firth's explanation of relationship among words is considered the initial theoretical framework. According to Firth (1957), "a collocation is a mode of meaning and the lexical meaning of a word is realised through multiple meanings on various levels" (Firth 1957:192). Other theoreticians define a collocation as an occurrence in which lexical units co-occur with one or more words (Halliday et al. 1964.:33; Ridout & Waldo-Clarke 1970.; Backlund 1973. 1976.; Seaton 1982.; Crystal 1985.:55; Cruse 1986.:40; Zhang 1993.:1). Collocations are very important in communication since their misuse can lead to misunderstanding. There is a great need for the analysis of collocations since they represent a connection between a word on the one hand and a text, on the other hand. However, there are not many dictionaries of collocations, especially those dealing with English for Specific Purposes (ESP).

That is why this research wants to investigate collocations in medical English trying to find out ways of simplifying their extraction from the corpus. So far, there have not been any studies in the field of medical English and most linguists were focused on the research of English for social sciences (e.g. Howarth 1998). Some researchers have dealt with collocations in natural sciences (Thomas 1993, Baker, Francis and Tognini-Bonelli 1993, Pearson 1998). The analysis of collocations requires many authentic texts and that is why a corpus is important. The first corpus used in the statistical analysis of a text was made in 1961 and had only 135,000 words. Two most famous corpora are – Brown’s corpus for American English and LOB (Lancaster – Oslo – Bergen) for British English. A corpus is particularly important in the collocation analysis, since it helps in obtaining data that can be statistically processed and is much more reliable way of checking collocations than the one that relies on native speakers (Sinclair 1991, Partington 1998, Hunston 2002, Krishnamurthy 2000). Thanks to computer technologies it started to be easier to make a corpus. All these are general English corpora without specialised vocabulary. A few lexicologists made a research on specialised corpora like Myers (1989), Kretzenbacher (1990), Banks (1994), Salager-Meyer (1992), Williams (1996), Dubois (1997) and Biber, Conrad and Reppen (1998). It is very important to have a corpus to study collocations and determine their frequency. Sinclair (1991), a pioneer of a corpus research thinks that in making a corpus it is very important to establish core vocabulary and wants to show that lexical repetition is very frequent in scientific terminology. Function/structure words are eliminated, thus leaving content/lexical words. The same principle was used in this research. The emphasis was made on upward collocations, i.e. those collocations where *a* is a collocate and *b* is a node (e.g. in a collocation ‘donate blood’, ‘donate’ is a collocate and ‘blood’ is a node). Each of the tools that were used had its function and they were all leading to the same aim – finding the most frequent nouns and their verb collocations that will be used in checking collocational competence of users of medical English and creating glossary of collocations that might prove to be useful to all the users of medical English.
The structure of the paper is as follows: first, the research programme is laid out. Then, methodology is explained which is followed by results and discussion. In the end, the conclusion is given.

Research questions

In order to achieve the aims of the study, the following research questions were formulated:

- What are the most frequent nouns in the corpus extracted by the corpus analysis tools?
- How can corpus analysis tools be used for collocation extraction?
- What is the Log Likelihood/Mutual Information of these collocations?

Methodology

First, the corpus was made from the online version of *Merck's Manual of Medical Information* (http://www.merckmanuals.com). The text of the book was turned into textual files that contain 1,065,181 words. There was a total of twenty five files that correspond to the chapters of the book: 1. Fundamentals; 2. Drugs; 3. Heart and Blood Disorders; 4. Lung and Airway Disorders; 5. Bone, Joint and Muscle Disorders; 6. Brain, Spinal Cord and Nerve Disorders; 7. Mental Health Disorders; 8. Mouth and Dental Disorders; 9. Digestive disorders; 10. Liver and Gallbladder Disorders; 11. Kidney and Urinary Tract Disorders; 12. Disorders of Nutrition and Metabolism; 13. Hormonal Disorders; 14. Blood Disorders; 15. Cancer; 16. Immune Disorders; 17. Infections; 18. Skin disorders; 19. Ear, Nose and Throat Disorders; 20. Eye Disorders; 21. Men's Health Issues; 22. Women's Health Issues; 23. Children's Health Issues; 24. Accidents and Injuries and 25. Special subjects. Next, the files were sent to the Faculty of Electrical Engineering in Zagreb where they were processed by TermeX tool in order to get the most frequent nouns. Simple Concordance tool was downloaded for processing the corpus to obtain concordances and find collocations. The last step was the use of Collocation Extract tool in order to find remaining collocations and determine their Log Likelihood and Mutual Information.

Tools and results

Three analysis tools are explained in more details to clarify their purpose. TermeX tool was used for extraction of the most frequent nouns in the medical corpus. Since it has limitations, i.e. it cannot extract verb collocations, Simple Concordance tool was used for this purpose. This tool is helpful for extracting concordance and finding adjacent collocations (the proximity is usually between one and three items). In order to find collocations in the range of 2 to 5 words, collocations extract was used and kit also established Mutual information and Log Likelihood. It is discussed below in more details.

TermeX

TermeX (http://takelab.fer.hr/TermeX/) (Seljan et al. 2009) is a tool for automatic collocation extraction and terminology lexical construction. The potential collocations are ranked by the strength of lexical associations; fourteen different lexical association measures are provided, based on Pointwise Mutual Information (PMI), Dice and Chi-square. The tool can extract collocations of up to length four. To this end, the standard bigram measures have been extended as proposed by Petrović et al. (2009). Moreover, TermeX tool enables the manual selection of candidate collocations to be included in terminology lexicon, the inspection of concordances of the extracted candidates, and efficient processing of multiple documents. However, its big disadvantage is that it does not extract verb collocations.

For this research, TermeX tool was used to provide most frequent nouns in the corpus and the number of their occurrences which can be seen in the Table 1. The Table 1. gives the list of ten most frequent nouns in the corpus[1]:

Table 1. The list of ten most frequent nouns in the corpus

Noun	Position	Number of occurrences
Blood	17.	6675
Symptoms	25.	4093
Treatment	38.	3059
Drugs	41.	2880
Heart	43.	2764
Disease	48.	2524
Pain	53.	2288
Infection	57.	2212
Skin	59.	2134
Body	61.	2008

The first column shows the noun, the second its position in the corpus and the third how many times it appears in the corpus. From the Table 1 it can be seen that the first most frequent noun, ‘blood’, appears on the 17th place and it appears 6675 times in the corpus. It is followed by ‘symptoms’, which appears on the 25th place and occurs 4093 times in the corpus. This list proved to be important later when the verb collocations of these nouns were determined.

Simple Concordance

Simple Concordance tool (http://www.textworld.com/scp) enables making a list of concordances that make it easier to find collocations. Each file was processed separately for easier text processing. For each word considered, the tool shows a list of concordances from corpus, known as *Key Words in Context (KWIC)*. Figure 1. gives an excerpt of concordances for the noun *blood.*

```
410          is also used to purify blood by removing harmful
410         or excessive numbers of blood cells or platelets in
410   . To be helpful for purifying blood, hemapheresis must
410    the undesirable substance or blood cell faster than the
411         that are used to purify blood are plasmapheresis and
411     , excess numbers of certain blood cells are removed.
411               (an excess of red blood cells), certain types of
411    leukemia (an excess of white blood cells), and
412      large fluid shifts between blood vessels and tissues that
412       and tissues that occur as blood is removed and returned
413   members or friends can donate blood specifically for one
413     the recipient's and donor's blood types and Rh factors are
413       , knowing who donated the blood is comforting, although
413   one from an unrelated person. Blood from a family member is
415    stem cells rather than whole blood. Prior to the donation
415     into the bloodstream. Whole blood is drawn from the donor,
```

[1] The list of other nouns is given in Miščin 2012.

```
415   a machine that separates the blood into its components
415    and returns the rest of the blood to the donor./DONATION
417      process of donating whole blood takes about 1 hour.
417      blood takes about 1 hour. Blood donors must be at least
417   in good health: their pulse, blood pressure, and
417            are measured, and a blood sample is tested to
418         a person from donating blood include hepatitis B or C
418       , poorly controlled high blood pressure, low blood
418       high blood pressure, low blood pressure, anemia, the
418     of hepatitis, and a recent blood transfusion./Generally,
419        are not allowed to give blood more than once every 56
419  practice of paying donors for blood has almost disappeared;
420    that would disqualify them./Blood Typing/Because
421     Typing/Because transfusing blood that does not match the
421      can be dangerous, donated blood is classified by type. A
421            by type. A person's blood type is determined by
421        proteins (Rh factor and blood group antigens A and B)
421   and B) on the surface of red blood cells./The four main
422     blood cells./The four main blood types are A, B, AB, and
422   and O, and for each type the blood is either Rh-positive or
422      , a person with O-negative blood has red blood cells that
422  with O-negative blood has red blood cells that lack both A
422    . A person with AB-positive blood has red blood cells that
422      AB-positive blood has red blood cells that have A and B
422        and the Rh factor. Some blood types are far more
422   than others. The most common blood types in the United
423  anyone can receive type O red blood cells; thus people with
423      ; thus people with type O blood are known as universal
423    donors. People with type AB blood can receive red blood
423  type AB blood can receive red blood cells from any blood
423       red blood cells from any blood type and are thus known
423   recipients. Recipients whose blood is Rh-negative must
423    is Rh-negative must receive blood from Rh-negative donors,
423   donors, but recipients whose blood is Rh-positive may
423     Rh-positive or Rh-negative blood./After a person is
424   is deemed eligible to donate blood, he sits in a reclining
424     the procedure is painless. Blood moves through the needle
424  bag. The actual collection of blood takes only about 10
```

Figure 1. The example of the part of concordances for the noun 'blood'

The collocations of the most frequent nouns were manually extracted. In that case, the most frequent combinations were two or three words away from the node, i.e., the central noun. Such an examination was necessary also to separate one-word nouns (e.g. blood) from the multi-word ones (e.g. blood cell, blood type, blood transfusion, etc.).

Collocation extract

The next tool is Collocation extract (http://collocation-extract.software.informer.com/3.0/), a tool for extracting collocations from a corpus. Collocation extract determines the lexical association word pairs based on the statistical measures of Log Likelihood (LL) and Mutual Information (MI). These measures help in determining collocation fixedness and measure the strength of lexical units and thus they help in distinguishing strong and weak collocations.

MI measures the statistical independence of words x and y by comparing their joint probability against the joint probability under the independence assumption. The use of MI for collocation extraction was first proposed by Church and Hanks (1990). The higher the MI, the higher the probability that words are lexically associated one with another. Log Likelihood is the probability ratio of the occurrence of one collocational component in the presence of another one and the probability that the same collocational component will occur without the other one. Higher LL denotes that the probability that two collocational components occur together is smaller. Collocation extract enables a user to determine the direction for searching collocations, the span, frequency, level of meaning and distance between two words. It was used for determining the connection between LL/MI and collocational competence, since it was expected that better knowledge of collocations was connected with higher LL/MI.

To process the corpus with Collocation extract, we proceed as follows. First, a corpus is chosen that should be in a plain text format and is put in ‘File-Save File List’. Then, statistical methods are chosen, in this case, Log-Likelihood and Mutual Information. Next, the span is chosen, which can be from 2 to 5. The number denotes the number of words in which collocations are searched for. For instance, if ‘2’ is chosen, the tool will look for two-word collocations (bigrams) and this was the number chosen here. Then, the direction for searching collocations is chosen. Since the upward collocations were looked for, the left side was chosen. After that, the minimum frequency of collocations was chosen. The lowest frequency of collocations was chosen, i.e. 1. Then, the statistical significance at the level ‘p<.005’, ‘p>.05’ or ‘all occurrences’ is selected. The option ‘all occurrences’ was selected and the maximum number of collocations that were extracted. The given value of ‘500’ was kept. In searching two-word collocations the distance between two words has to be determined. If ‘2’ is selected, two words are separated by one word. Since the ‘Simple Concordance’ tool had already extracted collocations with the distance of 3 or even 4, here, the option 1-6 was selected in order to determine other collocations with a greater distance between the members.[2] Table 2. gives the example of a part of the file obtained by such an analysis for the noun ‘diagnosis’.

As it can be seen from the Table 2., the collocate, i.e., the word before the researched noun, is in the first column. Articles, adjectives, nouns and verbs occur as collocates, but verbs were the subject of this research. The second column denotes the number of collocate occurrence, while in the third one the researched noun occurs, i.e. ‘diagnosis’. The fourth column mentions the number of its occurrence, which is 1712. After this, the frequency of occurrence of this collocation is shown (e.g. ‘confirm diagnosis’ occurs 133 times) and the last

[2] More detailed instructions on the use of the tool are mentioned in 'Help" of the tool itself.

column shows Log Likelihood. In this way, all the most frequent nouns were analysed and they were given in the table 3.

Table 2. The result of the analysis of 'Collocation Extract' tool for the noun 'diagnosis'

Word1	Freq1	Word2	Freq2	Freq12	ll
symptoms	4052	diagnosis	1712	439	4144.5531
the	63849	diagnosis	1712	597	2751.9236
and	29840	diagnosis	1712	446	2419.0548
confirm	185	diagnosis	1712	133	1850.5791
make	549	diagnosis	1712	66	621.725
confirms	31	diagnosis	1712	28	414.37192
makes	227	diagnosis	1712	39	396.75928
bases	23	diagnosis	1712	9	108.64861
a	24886	diagnosis	1712	58	102.06458
early	517	diagnosis	1712	15	97.100213
definitive	19	diagnosis	1712	7	83.43613
definite	6	diagnosis	1712	5	72.041094
making	242	diagnosis	1712	10	71.815283
establish	38	diagnosis	1712	6	59.819193
confirming	6	diagnosis	1712	4	54.318792
support	170	diagnosis	1712	7	50.208611
suspect	133	diagnosis	1712	6	44.141423
after	2315	diagnosis	1712	12	37.674508
establishing	12	diagnosis	1712	3	32.97642
considers	13	diagnosis	1712	3	32.428022
specific	486	diagnosis	1712	6	28.695834
suggest	98	diagnosis	1712	4	28.612737
preliminary	3	diagnosis	1712	2	27.157059
establishes	4	diagnosis	1712	2	25.431833
precise	47	diagnosis	1712	3	24.189261
prompt	65	diagnosis	1712	3	22.204354
suspects	71	diagnosis	1712	3	21.667737
accurate	71	diagnosis	1712	3	21.667737
suggests	79	diagnosis	1712	3	21.020867
the	63849	diagnosis	1712	8	19.74817
initial	100	diagnosis	1712	3	19.600239
screening	131	diagnosis	1712	3	17.985242

Table 3: The most frequent verb collocations in the corpus (occurring with the most frequent nouns)

COLLOCATION	TIMES IT APPEARS	LL	MI
receive a kidney	3	17.098098	7.5688532
aggravate the injury	1	8.2737904	12.467579
replace the hip	2	12.301151	10.280053
gain weight	18	210.23742	12.827786
establish diagnosis	6	59.819193	10.171127
tolerate pain	2	20.770593	6.3841196
provide relief	48	258.07393	11.071733
pose the risk	21	58.344398	9.9256855
loosen the secretion	5	59.901538	9.9880785
change the bandage	1	21.076902	16.327632
develop a bedsore	2	16.973625	12.581677
relieve pain	108	248.33414	8.4848646
treat the infection	35	44.184953	38.331006
regain consciousness	1	44.825846	12.10363
induce vomiting	5	76.444688	10.512782
produce pain	3	51.985936	5.5569562
suppress inflammation	3	35.032721	9.0880331
undergo dialysis	2	122.05025	12.285744
detect a lump	1	14.268766	9.6039706
impair memory	2	23.424866	9.8576507
abort headaches	3	39.895228	9.8576507
relieve nausea	4	30.481541	6.9119118
tolerate a drug	1	10.62278	4.4610449
catch a cold	1	16.128507	9.1082614
detect a cancer	10	65.871482	5.9893827
cleanse the wound	2	20.29615	12.683775
transmit a disease	17	139.06414	7.4514984
get/develop symptoms	11/47	22.681104/ 201.61944	7.9952898/ 9.9281756
identify antibodies	1	24.756581	9.291458
relieve anxiety	5	9.0133363	7.9225788
precipitate the attack	1	13.334497	10.904165
suppress a cough	4	70.689533	10.66052
cause discomfort	16	55.418659	5.603248
trigger diseases	1	11.883876	9.9288879

Table 3[3]. shows the most frequent upward verb collocations in the corpus which occur with the most frequent nouns, the frequency of their occurrence, their Log Likelihood and Mutual Information. It can be seen that the most frequent collocation is 'relieve pain', which occurs 108 times and also has quite a high Log Likelihood (248.33414), but not so high Mutual Information (8.4848646). The next one is 'provide relief', which occurs 48 times with also high Log Likeli-

[3] This is only a part of the table. The whole table can be seen in Miščin 2012.

hood (258.07393) and low Mutual Information (11.071733). They are followed by 'treat the infection' (35 times) which has a lower LL (44.184953) and a bit higher MI (38.331006).

Conclusion

Processing of any corpus would be a difficult task without computer tools, especially when dealing with collocations. This paper showed the use of three different corpus analysis tools – TermeX, Simple Concordance, and Collocation Extract in processing of the medical corpus. TermeX was useful for making a list of the most frequent nouns and analysing a list of their frequencies. It was shown that the most frequent noun is blood which occurs 6675 times. Simple Concordance proved to be useful in making a list of concordances and finding collocations. Collocation Extract helped in extracting collocations occurring with the most frequent nouns. It was established that the most frequent collocations were 'relieve pain', 'provide relief' and 'treat the infection'. The next step would be to test collocational competence of users of medical English and to establish the connection between the competence and Log Likelihood/Mutual Information. The aim of all these analyses was to analyse the results of automatic extraction results. The future research would aim to consider collocations which would be the most useful for users of medical English and to make a dictionary of medical English collocations.

References

Backlund, U. The collocation of adverbs of degree in English. Doctoral Dissertation, Uppsala University, Uppsala, Sweden, 1973.

Baker, M.; Francis, G. and Tognini-Bonelli, E. (eds.). Text and Technology. Amsterdam: John Benjamins, 1993.

Banks, D. 'Clause Organisation in the Scientific Journal Article'. *Alsed-Lsp Newsletter* (1994). Vol.17/2.4-16

Biber, D.; Conrad, S. and Reppen, R. Corpus Linguistics: Investigating Language Structure and Use. Cambridge: Cambridge University Press, 1998.

Church, Kenneth Ward; Hanks, Patrick. Word Association Norms, Mutual Information and Lexicography. In *Proceedings of 27th ACL,* 16(1) (1990), 22-29

Cruse, D.A. Lexical Semantics, Cambridge: Cambridge University Press. 1986

Crystal, David. A dictionary of linguistics and phonetics. Oxford: Basil Blackwell Ltd., 1985.

Delač, Davor; Krleža, Zoran; Dalbelo Bašić, Bojana; Šnajder, Jan; Šarić, Frane. TermeX: "A Tool for Collocation Extraction. Lecture Notes" in Computer Science" (Computational Linguistics and Intelligent Text Processing). 5449 (2009); 149-157

Dubois, B.L. The Biomedical Discussion Section in Context. London: Ablex Publishing Corporation, 1997

Dunning, Ted. Accurate Method for the Statistic of Surprise and Coincidence. In *Computational Linguistics,* (1993), 61-74

Firth, J.R. Papers in Linguistics 1934-1951. Oxford: oxford University Press, 1957.

Halliday, M.A.K. et al. The linguistic sciences and language teaching. London: Longman, 1964.

Howarth, P. "The Phraseology of Learners' Academic Writing' in Cowie, A.P. (ed.). Phraseology: Theory, Analysis and Applications. Oxford: Oxford University Press, 1998.

Hunston, Susan. Corpora in Applied Linguistics. Cambridge: Cambridge University Press, 2002.

Kretzenbacher, H.L. Rekapitulation: Textstrategien der Zusammenfassung von Wissenschaftlichen Fachtexten. Tübingen: Gunter Narr Verlag, 1990.
Krishnamurthy, Ramesh. “Collocation from silly ass to lexical sets” in Heffer, C; Sauntson, H. and Fox, G. (eds.) Words in Context: A tribute to John Sinclair on his Retirement. Birmingham: University of Birmingham, 2000.
Miščin, Evelina. Unpublished doctoral thesis ‘Glagolske kolokacije u engleskome jeziku’. Osijek, 2012.
Myers, G. ‘The Pragmatics of Politeness in Scientific ARticles.’ In *Applied Linguistics* (1989). Vol. 10/1:1-35.
Partington, Alan. Patterns and Meanings. Amsterdam: John Benjamins, 1998.
Pearson, J. Terms in Context. Amsterdam: John Benjamins, 1998.
Petrović, Saša, Šnajder, Jan and Dalbelo Bašić, Bojana. “Extending lexical association measures for collocation extraction.” *Computer Speech & Language* 24.2 (2010): 383:394.
Ridout, R. & Waldo-Clarke, D. A reference book of English. London: Macmillan, 1970.
Salager-Meyer, F. ‘A Text-Typer and Move Analysis Study of Verb Tense and Modality Distribution in Medical English Abstracts.’ In *English for Specific Purposes* (1992) Vol.9: 145-159.
Seaton, B. A handbook of English language teaching terms and practice. London: The Macmillan Press Ltd., 1982.
Seljan, Sanja, et al. “Comparative Analysis of Automatic Term and Collocation Extraction.” *The Future of Information Sciences: INFuture 2009-Digital Resources and Knowledge Sharing/ Stančić, H. (2009): 219-228.*
Seljan, Sanja; Gašpar, Angelina. First Steps in Term and Collocation Extraction from English-Croatian Corpus // Proceedings of 8th International Conference on Terminology and Artificial Intelligence. Toulouse, France, 2009. http://www.irit.fr/TIA09/thekey/posters/seljan.pdf
Sinclair, John. Corpus, Concordance, Collocation. Oxford: Oxford University Press, 1991.
Thomas, P. ‘Choosing Headwords from LSP Collocations for Entry Into A Terminology Data Bank (Term Bank).’ In Sonneveld, H.B. and Loening K.L. (eds.) 1993: 46-68.
Williams, I.A. ‘A Contextual Study of Lexical Verbs in Two Types of Medical Research Article.’ In *English forSpecific Purposes.* (1996). Vol. 15/3:175-198.
Zhang, X. English collocations and their effect on the writing of native and non-native college freshmen. Unpublished Ph.D. thesis, Indiana University of Pennsylvania, 1993.

Web links

http://www.merckmanuals.com
http://takelab.fer.hr/TermeX/
http://www.textworld.com/scp
http://collocation-extract.software.informer.com/3.0/

DIGITAL CURATION

Crowdsourcing Digital Cultural Heritage

Goran Zlodi
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
gzlodi@ffzg.hr

Tomislav Ivanjko
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
tivanjko@ffzg.hr

Summary

With the turn towards the digital age, a growing number of institutions in the GLAM (Galleries, Archives, Museums and Libraries) sector started to identify a need for digitising their different collections placing them online with goals to preserve and exhibit them in the digital environment. After the initial efforts to develop policies, methodologies and best practices in transferring the collections into the online environment, researchers and practitioners have started to investigate possibilities of communicating those digitised collections with the public and seizing the opportunities that arise from digitisation. One of the approaches that cultural heritage institutions started to explore in order to involve the general public in their activities on the Web is crowdsourcing - taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined community in the form of an open call. In the heritage sector this means inviting members of the public, ("the crowd"), to tag and classify, transcribe, organize, and otherwise add value to digital cultural heritage collection content. In this paper we provide an overview of approaches in using the collective intelligence in the cultural heritage domain. Key terms, concepts and corresponding case studies are discussed, providing the framework for crowdsourcing projects within the heritage sector.

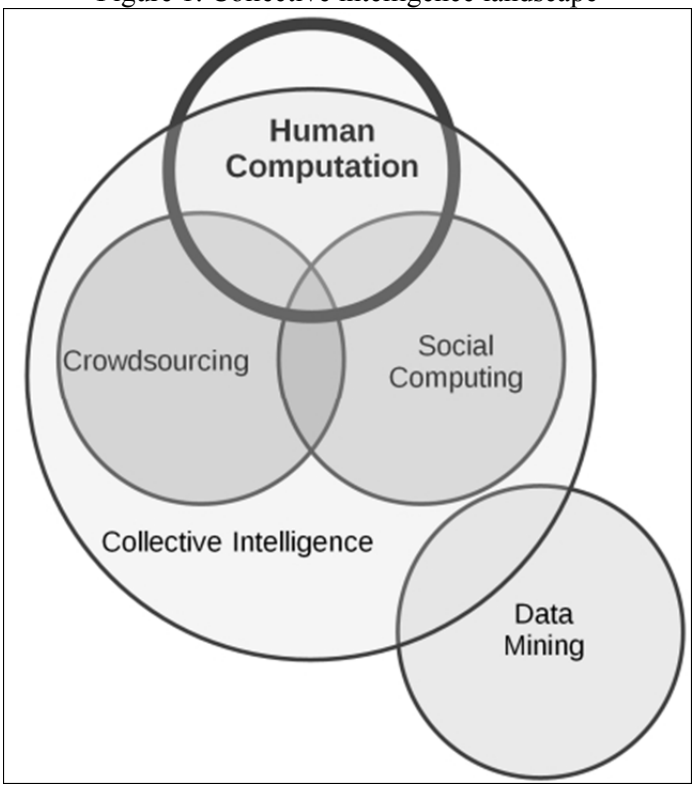
Key words: crowdsourcing, heritage, Web 2.0, archives, libraries, museums

Harnessing collective intelligence in the heritage sector – defining the field

When the notion of crowdsourcing is explored within the scientific literature related to the field, a number of related terms emerge. Crowdsourcing is often related with terms such as “collective intelligence”, “wisdom of the crowds” “human computation”, “social computing”, etc.

Figure 1 presents an attempt to classify the landscape of various systems harnessing the collective intelligence where the relations between the main concepts in the field are presented (Bederson and Quinn, 2011.)

Figure 1: Collective intelligence landscape



Authors identify three main notions that comprise the field of collective intelligence: human computation, crowdsourcing and social computing. All three notions fall in the field of collective intelligence, by having the same prerequisite for successful implementation – they all depend on a group of participants (Bederson and Quinn, 2011).

Within this framework, the most interesting is the differentiation between human computation and crowdsourcing. The modern usage of the term human computation is largely credited to the title of the thesis from of Luis von Ahn

and the related papers (von Ahn, 2005; Law and von Ahn, 2009; von Ahn and Dabbish, 2008). In his thesis (von Ahn, 2005), von Ahn defines the term as "…a paradigm for utilizing human processing power to solve problems that computers cannot yet solve." The main idea behind human computation is using the collective intelligence of users to solve problems that are hard or still impossible to do by using computer programs or algorithms. One of the most common examples of a human computation system is ReCAPTCHA (recaptcha.net), used for transcribing scanned texts for which OCR is not very effective. It takes advantage of the need for CAPTCHAs, the distorted images of text that are used by websites to prevent access by automated programs (von Ahn et al., 2008).
On the other hand, term crowdsourcing is derived from the word outsourcing where a job traditionally performed by a designated agent (usually an employee) is outsourced it to an undefined, generally large group of people in the form of an open call (Howe, 2008). As Bederson and Quinn (2009) summarize, the difference between crowdsourcing and human computation is that "…Whereas human computation replaces computers with humans, crowdsourcing replaces traditional human workers with members of the public." Owens (2012) also notes that crucial difference between those two perspectives (using slightly different terminology), considering human computation and the wisdom of crowds as opposing polls of crowdsourcing activity and provides an overview of key differences between them (Table 1)

Table 1: Key differences between human computation and wisdom of crowds

	Human Computation	**Wisdom of Crowds**
Tools	Sophisticated	Simple
Task Nature	Highly structured	Open ended
Time Commitment	Quick & Discrete	Long & Ongoing
Social Interaction	Minimal	Extensive Community Building
Rules	Technically Implemented	Socially Negotiated

When thinking of human computation, one should imagine the example of the ESP game (link) which Owens (2012) describes as "a sophisticated little tool that prompts us to engage in a highly structured task for a very brief period of time...with almost no time commitment...practically no social interaction...and the rules of the game are strictly moderated by the technical system." On the other hand Wikipedia being the example of wisdom of the crowds: "While the tool is very simple the nature of our task is huge and open-ended...it's open-ended nature invites much more long-term commitment...an extensive community building process..." (Owens, 2012).
Following the rise of the Web 2.0, heritage institutions quickly realized the potential of the fundamental ideas underlying Web 2.0 - successful network applications are systems for harnessing collective intelligence (O'Reilly and Batelle, 2009). Many libraries, institutions or archives started using different Web 2.0 tools, such as blogs or social networks or to extend and enhance their communi-

cation with their users. According to ICOM, the museum “...acquires, conserves, researches, communicates and exhibits the tangible and intangible heritage of humanity and its environment for the purposes of education, study and enjoyment” (ICOM, 2013). That can be said for basically all institutions within the heritage sector so the idea of crowdsourcing fitted right in with the goals of researching and communicating heritage for different purposes and users. After the initial projects explored the field, some researchers found the term “crowdsourcing” isn’t really appropriate for many cultural heritage projects. Both the appropriateness of the term “crowd” or the term “sourcing” is questioned, because such projects in the heritage sector mostly don’t involve large and massive crowds and have very little to do with outsourcing labor. (Owens, 2013). One of the terms that can offer new perspectives is “nichesourcing”, a “...specific type of crowdsourcing where complex tasks are distributed amongst a small crowd of amateur experts ... rather than the “faceless” crowd” (de Boer et al., 2012). As Owens (2012) argues, projects within the cultural heritage sector are mostly about just that: “...inviting participation from interested and engaged members of the public... [they] continue a long standing tradition of volunteerism and involvement of citizens in the creation and continued development of public goods (Owens, 2013). Following these arguments, we can see that the original Web 2.0 idea of harnessing collective intelligence found its application in the cultural heritage but denoting a specific field of application.

The initial idea offered a broad perspective with the idea that a large group of people can create a collective work whose value far exceeds that provided by any of the individual participants (O’Reilly and Batelle, 2009). This is closely related to the idea of social computing, and opened a door for different Web 2.0 applications such as Facebook, Youtube or Wikipedia where the collective power of the users was harnessed in its full strength. Along with that the field of human computation investigated how using the collective intelligence of users can solve problems that are hard or still impossible to do by using computer programs or algorithms (von Ahn, 2005). This produced many great applications such as the ESP game or reCaptcha where it was clearly shown how many users can overcome serious computing problems, such as image labelling on the Web. The third incentive came in the coining of the idea of crowdsourcing – outsourcing a job usually carried out by employees to the public (Howe, 2008). One example being the annotation of large collections of digitized materials that require a lot more human effort than the institution employees can handle. Although these goals can be accomplished by crowdsourcing, because of the need for specific knowledge on the subject matter often the “public” is not the right target if the quality of metadata gathered is vital. One approach can be found in the idea of “nichesourcing”, a further development of the crowdsourcing idea focused on solving different complex knowledge-intensive tasks and providing quality results by involving amateur experts instead of the general “faceless crowd” (de Boer et al., 2012).

Taking all these perspectives in consideration we can see that the field of harnessing collective intelligence in the cultural heritage sector is not only based on the idea to *use* the public, but to engage them to contribute, collaborate and co-create (Bonney et al., 2008). It is not only about getting things done (e.g. describe or transcribe certain materials) but to communicate the collection to the users by shifting their focus from consuming digital collections to collaborating in its development. This most commonly leads to a group of amateur experts that have joined the project because they care about the cause and have intrinsic motivations to participate. Both of these aspects underline the paradox of using the term "crowdsourcing" when describing cultural heritage projects since they, in most cases, include engaged amateurs (not the "crowd") that are intrinsically motivated and don't consider their work as labour (thus nothing is really „outsourced"). Since the term is already in use, there is a need for refinement and distinction when using the term "crowdsourcing" in the cultural heritage domain by including all the perspectives from human computation to nichesourcing as parts of the field itself.

Setting the framework for applying crowdsourcing in the cultural heritage domain

Growing number of institutions in the GLAM (Galleries, Archives, Museums and Libraries) sector started to investigate possibilities of communicating their digitised collections with the public and seizing the opportunities that arise from digitisation by applying different crowdsourcing approaches. For that reason, it is important to classify the different types of crowdsourcing in the GLAM domain so potential new projects have an overview of opportunities and challenges in the domain. In this chapter we discuss the existing attempts to classify the field of crowdsourcing with special focus on the digital heritage sector.

One of the attempts to provide a general overview of the field from the *service perspective* is "Crowdsourcing Industry Landscape", an infographic presenting crowdsourcing taxonomy in order to provide a framework for the industry. The infographic is constantly revised, with the latest version reflecting the third generation of the taxonomy, categorising the field in six different areas:

1. crowdfunding – financial contributions from online investors, sponsors or donors to found for-profit or non-profit initiatives or enterprises (e.g. kickstarter.com, gofundme.com)
2. crowd creativity – tapping of creative talent pools to design and develop original art, media or content (e.g. istockphoto.com, minted.c
3. distributed knowledge – development of knowledge assets or information resources from a distributed pool of contributors (e.g. GalaxyZoo, openbuildings.com)
4. cloud labour – leveraging of a distributed virtual labour pool available on-demand to fulfil a range of tasks from simple to complex(e.g. AmazonMechanicalTurk, tagasauris.com)

5. open innovation – use of sources outside of the entity or group to generate, develop and implement ideas (e.g. challengepost.com, innocentive.com)
6. tools – applications, platforms and tools that support collaboration, communication and sharing among distributed groups of people (e.g. socialvibe.com, bigdoor.com)

When approaching these categories from a heritage perspective, the category of distributed knowledge stands out as a category where crowdsourcing digital cultural heritage fits right in. As we mentioned earlier, when defining the field, the idea is to engage motivated users to contribute, collaborate and co-create, which is closely related to the distributed knowledge category description and the projects listed there. Although these categories are mainly aimed at the industry and applications, this overview is a good place to start because it lists a large number of actual projects and ideas in the field and is very informative on the possible approaches when thinking about implementing crowdsourcing in the heritage sector. Since this categorisation is very broad there is a need to look at he attempts to classify the specific domain of crowdsourcing in the cultural heritage domain. Based on the approaches in the literature the field can be categorized through three aspects:

1. participation – categorising crowdsourcing projects based on the level of user engagement
2. activities – categorising crowdsourcing projects on the types of activities undertaken
3. crowdsourcing initiatives – categorising crowdsourcing projects based on the tangible outcomes

These aspects represent the three main questions every institution should answer when considering crowdsourcing their collection, namely who are the intended users, what activities should be implemented and what are the tangible outcomes.

In the book *The Participatory Museum* (Simone, chapter 5) author list four models based on public participation in cultural institutions: (1) *contributory* projects, where visitors are solicited to provide limited and specified objects, actions, or ideas to an institutionally controlled process; (2) *collaborative* projects, where visitors are invited to serve as active partners in the creation of institutional projects that are originated and ultimately controlled by the institution; (3) *co-creative* projects, community members work together with institutional staff members from the beginning to define the project’s goals and to generate the program or exhibit based on community interests and (4) *hosted* project in which the institution turns over a portion of its facilities and/or resources to present programs developed and implemented by public groups or casual visitors.

This classification looks at the field through different model of participation, providing a framework on how an institution could engage its users. As the au-

thor puts it: "No one model is better than the others. Nor should they even be seen as progressive steps towards a model of "maximal participation"...The differences among participatory project types are highly correlated with the amount of ownership, control of process, and creative output given to institutional staff members and visitors." As a starting point for every institution, a chart displaying the fundamental characteristics of each model is provided (Simone, 2010).

Complementing the models of participation, a decision on *crowdsourcing activities* is another important aspect of the process. Ridge (2011) looks at the types of various activities which can be applied to digitized objects. Although her work primarily deals with the context of things people can do with museums and games to improve museum collections, this categorisation can be applied to general crowdsourcing activities:

- tagging – applying unstructured labels to individual objects
- debunking – flagging content for review and/or researching and providing corrections
- linking – linking objects with other objects, objects to subject authorities, objects to related media or websites;
- categorising – applying structured labels to a group of objects, collecting sets of objects or guessing the label for or relationship between presented set of objects
- stating preferences – choosing between two objects or voting on or 'liking' content
- recording a personal story – contextualising details by providing subjective oral histories or eyewitness accounts
- creative responses – writing an interesting fake history for a known object or purpose of a mystery objects.

This overview briefly summarizes what types of crowdsourcing activities can be implemented when dealing with objects in the cultural heritage sector.

Table 2: Types of crowdsourcing initiatives

Crowdsourcing type	**Sort definition**
Correction and Transcription Task	Inviting users to correct and/or transcribe outputs of digitisation processes.
Contextualisation	Adding contextual knowledge to objects, e.g. by telling stories or writing articles/wiki pages with contextual data.
Complementing Collection	Active pursuit of additional objects to be included in a (Web)exhibit or collection.
Classification	Gathering descriptive metadata related to objects in collection. Social tagging is a well-known example.
Co-curation	Using inspiration/expertise of non-professional curators to create (Web)exhibits.
Crowdfunding	Collective cooperation of people who pool their money and other resources together to support efforts initiated by others.

The final aspect of the decision on implementing crowdsourcing is looking at the field from the perspective of tangible outcomes, i.e. how can different crowdsourcing types contribute to the working practices and what can different initiatives offer as real outcomes. Oomen and Arroyo (2011) list the classification of crowdsourcing activities based on their outcomes, a result of a study gathering examples of crowdsourcing initiatives around the globe (Table 2).

Conclusion

There are many advantages in implementing crowdsourcing services and projects within cultural heritage institutions. Users tagging, annotating and adding contextual knowledge results in easier managing of huge collections and demonstrates obvious benefits for galleries, archives, museums and libraries. Important added value for institutions is strengthening the relations with end users and consequently getting more precise insight in user’s needs. On the other side, persons involved in crowdsourcing also find value in contributing to cultural heritage research and enrichment of their cultural identity. In this new way of partnership in digital environment it is necessary to continuously improve mechanisms of collaboration to achieve desired level of trustworthiness and quality of added content. Such collaboration of heritage professionals and involved end-users will as a result bring data sets of different origins to exist in the same databases. However, in presenting of content through public web services, origin of data will have to be clearly labeled so that user created data is easily differentiated from data created by heritage professionals.

References

Bonney, R. et al. Public participation in scientific research: defining the field and assessing its potential for informal science education. A CAISE Inquiry Group report. Washington, D.C., Center for Advancement of Informal Science Education (CAISE), 2009.

Crowdsourcing industry landscape. 2011. http://www.crowdsourcing.org/document/december-2011-crowdsourcing-industry-landscape-download-version/9664 (2013/09/11)

de Boer, V. et al. Nichesourcing: harnessing the power of crowds of experts. // Knowledge Engineering and Knowledge Management. Proceedings of the 18th International Conference, EKAW 2012. Berlin Heidelberg, Springer, 2012, 16-20.

Howe, J. Crowdsourcing: a definition. 2010. http://crowdsourcing.typepad.com (2013/09/11)

Howe, J. Crowdsourcing: why the power of the crowd is driving the future of business. Crown, 2008.

Howe, J. The rise of crowdsourcing. // Wired magazine. 14 (2006), 6; 1-4.

Law, E.; von Ahn, L. Input-agreement: a new mechanism for collecting data using human computation games. // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). New York: ACM, 2009. 1197-1206.

Oomen, J.; Aroyo, L. Crowdsourcing in the cultural heritage domain: opportunities and challenges. // *Proceedings of the 5th International Conference on Communities and Technologies* (C&T '11). New York: ACM, 2011, 138-149.

O'Reilly, T.; Battelle, J. Web squared: Web 2.0 five years on. O'Reilly Media, 2009. http://gossgrove.com/sites/default/files/web2009_websquared-whitepaper.pdf

Owens, T. Digital cultural heritage and the crowd. // Curator: The Museum Journal. 56 (2013), 1; 121-130.

Owens, T. Human Computation and Wisdom of Crowds in Cultural Heritage. 2012. http://www.trevorowens.org/2012/06/human-computation-and-wisdom-of-crowds-in-cultural-heritage/ (2013/09/11)

Quinn, A. J.; Bederson, B. B. Human computation: a survey and taxonomy of a growing field . // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’11). New York: ACM, 2011, 1403-1412.

Ridge, M. Playing with difficult objects – game designs to improve museum collections. // Museums and the Web 2011: Proceedings / Trant, J.; Bearman, D. (eds). Toronto: Archives & Museum Informatics. 2011. http://conference.archimuse.com/mw2011/papers/playing_with_difficult_objects_game_designs_improve_museum_collections

Simon, N. The Participatory Museum. Santa Cruz, California: Museum 2.0, 2010.

von Ahn, L. et al. ReCAPTCHA: human-based character recognition via web security measures. // Science. 321 (2008), 5895; 1465-1468.

von Ahn, L. Games with a Purpose. // Computer. 39 (2006), 6; 92-96.

von Ahn, L. Human Computation. Doctoral Thesis. Carnegie Mellon University, 2005.

von Ahn, L.; Dabbish, L. Designing games with a purpose. // Communications of the ACM. 51 (2008), 8; 58-67.

von Ahn, L.; Dabbish, L. Labeling images with a computer game. // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04). New York: ACM, 2004, 319-326.

von Ahn, L.; Kedia, M.; Blum, M. Verbosity: a game for collecting common-sense facts. // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06) / Grinter, R. et al. (eds.). New York: ACM, 2006., 75-78.

Bringing Circus History into the Digital World

Maureen Brunsdale
Milner Library, Illinois State University
Normal, IL 61790-8900, U.S.A.
mbrunsdale@ilstu.edu

Summary

No library or museum in the world was curating a growing online exhibit of circus images when Milner Library at Illinois State University began such an exhibit in the spring of 2010. Images were first hosted locally and then nationally, but each time the platform used was the Online Computer Library Center's (OCLC) collection management software, CONTENTdm. To date, the collection has grown to include over 5000 digitized preserved images culled from both Kodachrome slides and black and white 35mm negatives. This paper will demonstrate how CONTENTdm enhances the digital curation and preservation of the online exhibit known as "Passion for Circus," ensuring this truly interdisciplinary art form's existence well into the future.

Keywords: Digital curation, online exhibit, CONTENTdm, circus

Introduction

In 2008, Taschen Press released a magnificent book called *Circus, 1870-1950*. Weighing 15.5 pounds, this gigantic book was printed in English, French and German and was illustrated by more than 1,000 images. A reviewer for the *New York Times Book Review* commented, "This is a gee-whiz spectacle of a book, a three-ring extravaganza as bright as a pinball machine." The book's reception was equally vivid, garnering four more printings in the five years since it was first published. The central feature of the book are its many photographs, roughly one-third of which came from the photographs of Sverre O. Braathen, a significant portion of the Circus & Allied Arts Collection found within Special Collections at Milner Library, Illinois State University, Normal, Illinois.

The publication of the first edition of this book coincided with an annual convention of the Circus Historical Society, held in Normal. The coalescence of all these events gave the idea to digitize of the circus photography of Sverre O. Braathen the momentum it had never previously had. The software chosen to make this rich primary source material available on the internet was Online Computer Library Center's (OCLC) product named CONTENTdm. This software has ably preserved the images from a significant cultural and demographic entity, one that in many ways defines interdisciplinary studies. (Another sentence, maybe, to that effect?...something like: From the study of the artistic

composition of the photographs, to the discovery of socioeconomic indicators revealed in them, to the analysis of gender roles in mid-twentieth century performance arts, these digitized images offer rich insights from a spectrum of academic and cultural perspectives—or something.)

How the software works – a user’s perspective

OCLC’s digital collection management software, CONTENTdm, provides many benefits to the user. A look at “Passion for Circus” – the name for the aforementioned collection of photographs – highlights many of those benefits. If someone wanted to do some research on this collection, he or she might open a web browser and search for the collection title, Passion for Circus, using a search engine such as Google. That sort of scenario reveals the primary benefit of using CONTENTdm: the collection is easily retrieved—across many platforms and search engines. The front page for “Passion for Circus” would be displayed, having the same look and feel as the rest of the webpages for Milner Library, maintaining the branding of our institutional web presence with the seal of Illinois State University displayed next to the name of our library, followed by the title of this digitized collection. Under the red banner which helps to serve as the university brand, is a simple search box, wherein a user could search for images related to his or her research need. One of the Braathen images fills the middle of the screen, then toward the bottom of the page, the user can choose to search the collection by a variety of means: Circus, Performer, Stage Name, Location and Subject. The right-most side of the screen offers a link users can follow to get a Really Simple Syndication (RSS) feed, providing an alert when any new content is added to the collection. Finally, thumbnails of the most recently added images to the project are displayed at the bottom of the page. All of these components illustrate some of the benefits of using CONTENTdm.

An example helps to illustrate the functionality. Recently, a researcher was looking for images of the air conditioning units which the Ringling Brothers & Barnum and Bailey Circus used in the middle of the last century. By typing “air conditioning” into the search statement box, 37 results were produced. Upon first glance, many – like the “Old King Cole Spectacle Entering the Big Top” - do not appear to relate to the concept.

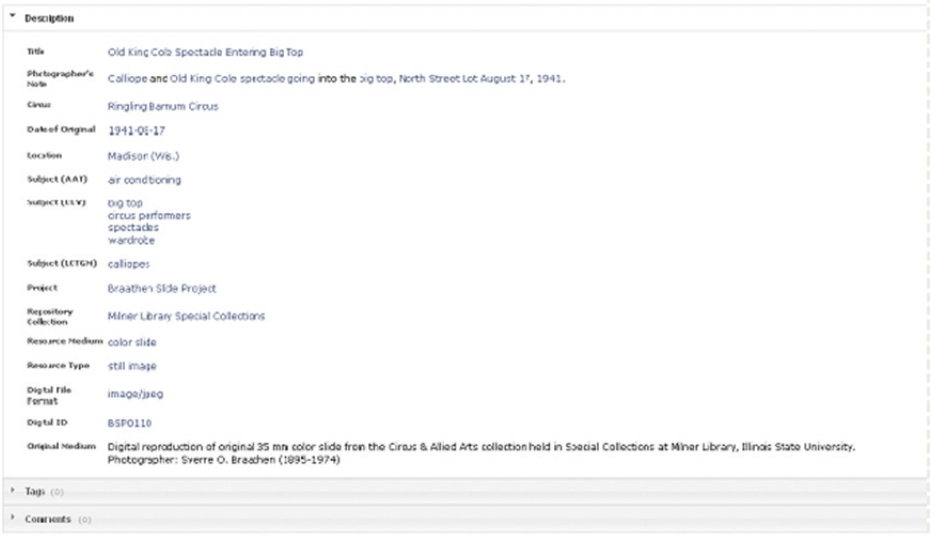
This gives the opportunity to display one of the greatest technical features of CONTENTdm: the ability to “zoom in” on the photo. Doing that while moving the mouse to highlight the top third of the image shows a flattened tube with the upside-down letters that read, “AIR COND.” These first few letters alert us to the fact that this flattened tube was part of the air conditioning apparatus used by that circus.

Photo 1

Source: http://digital.library.illinoisstate.edu/cdm/singleitem/collection/circus/id/535/rec/12

But what year was this tube used? A look at the record which accompanies the image gives us the data we need to ascertain this vital piece of information - and so much more.

Screen shot 1



Description

Title	Old King Cole Spectacle Entering Big Top
Photographer's Note	Calliope and Old King Cole spectacle going into the big top, North Street Lot August 17, 1941.
Circus	Ringling Barnum Circus
Date of Original	1941-08-17
Location	Madison (Wis.)
Subject (AAT)	air conditioning
Subject (LCSH)	big top circus performers spectacles wardrobe
Subject (LCTGM)	calliopes
Project	Braathen Slide Project
Repository Collection	Milner Library Special Collections
Resource Medium	color slide
Resource Type	still image
Digital File Format	image/jpeg
Digital ID	BSP0110
Original Medium	Digital reproduction of original 35 mm color slide from the Circus & Allied Arts collection held in Special Collections at Milner Library, Illinois State University. Photographer: Sverre O. Braathen (1895-1974)

Tags (0)

Comments (0)

The record for the photo above detailing all the metadata for it.
http://digital.library.illinoisstate.edu/cdm/singleitem/collection/circus/id/535/rec/12

The date - August 17, 1941 – is provided along with the title of the image, the location where it was filmed (Madison, Wisconsin), the subject fields (air conditioning, big top, circus performers, spectacles, wardrobe), and all the ownership information can all be found in that record. Librarians call these things the metadata, but users likely refer to them as the details about each item. Regardless of what they are called, they are the priceless specifics which help define each image. They also provide the context for them. Finally, links provided at the bottom of the screen allow users to tag items, add commentary, and/or offer corrections if mistakes were made in identifying people or objects.

Additional Benefits

Beyond ease of use and searchability across many platforms and search engines, CONTENTdm provides additional benefits. For instance, the software delivers the means for a library to digitally store and thus preserve unique and valuable images. It also is customizable on the back-end which in our case means that the initial search fields can be designed to reflect the unique vocabulary of the circus.

Internet users are a diverse lot. Many enjoy using social media; CONTENTdm allows for image-sharing via Facebook, Twitter, Flickr and more. Those searching the web also sometimes desire to purchase or use for research images they find there. An exciting feature has just been released by OCLC and their partner, Atlas Systems which addresses this issue. Atlas Systems’ Aeon software works with CONTENTdm to allow users to do two things: to request that materials to be pulled for them before physically arriving at a Special Collections depository, and it also enables people to not only identify which images they would like to have reproduced for publication or personal use, but also to pay for the reproduction services online via a credit card. The press release for this feature was disseminated in January, 2013. As yet, Milner Library has not implemented this service.

Conclusion

In the span of a few short years Milner Library has introduced the world to an online photo exhibit of circus images that documents American circus shows from several decades of the twentieth century. This has been accomplished by first digitizing the slides, negatives, and photographs from the Circus & Allied Arts Collection found within the Special Collections department, and then uploading them into OCLC’s online collection management software, CONTENTdm. Images from this collection can be found in wildly diverse websites, from camera and film fan sites, to sites whose supporters love Jeeps, to model builders who want detailed images of Caterpillar tractors from the 1940s. Just like the circus itself, the study of digitized circus images attracts diverse crowds from academics, to independent scholars, as well as circus fans. The digitization of this collection of photographs ensures that more people are able to access these wonderful images, in effect bringing a page from American cultural history to life once again. We look forward to expanding the offerings of images and services in the future so that the more people may be able to access these wonderful images from days gone by, in effect bringing a segment of American cultural history to life once again.

NEW CHALLENGES IN INTERDISCIPLINARY EDUCATION

Original scientific paper

Interactive Application for Learning the Latin Language

Matea Leko, Sanja Kišiček, Josip Knežević,
Petar Krešimir Marić, Ivana Martinović, Elvis Vusić
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
mleko@ffzg.hr, sanja.kisicek@gmail.com, jknezevi@ffzg.hr,
pkmaric@ffzg.hr, imartino2@ffzg.hr, evusic@ffzg.hr

Summary

Computers in education are an integral part of the teaching process, but the use of them mostly depends on the competence and willingness of individual teachers, as well as the availability of adequate hardware and software. When it comes to learning a language, computer assisted language learning software mostly covers languages such as English, or German, while other, less popular languages still do not have a good software foundation.
The interactive application for learning the Latin language is an example of software that seeks to facilitate students studying Latin within all grammar schools in Croatia. The Latin language, although very often called a dead language, is still a part of the culture in general, and important in fields such as medicine, law, agriculture, etc. Results obtained by our research confirm that students are willing to use such software and that the increase in student motivation and interest are proportional to the final test results.

Key words: computer assisted language learning, the Latin language, interactive software

Introduction

Computers significantly affect the concept of modern teaching, whether used by a teacher as a classroom tool, or independently by students. Teachers use computers in classrooms for different purposes. The goal is always to improve student learning and not only make the learning material easily available to students. Computers are frequently used in order to accommodate different learning needs as well as to motive the learners in general. However, the use of computers only is not enough to maintain student attention at a high level during a lesson. Students have embraced computers as an integral part of teaching and therefore expect it to be applied in interesting ways, very often through fun and playing games.

Since many teachers lack the competence or resources to create computer applications for learning, teachers mostly rely on the existing ones. The usefulness of applying computer assisted learning (CAL) is highlighted in the area of foreign language learning which requires constant practice and repetition. CAL enables students to make decisions on their own about the level of their knowledge, significance of the subject matter and their learning pace (Parsons and Oja, 2010). Within the course *Computer Assisted Teaching* at the Faculty of Humanities and Social Sciences in Zagreb, we designed and created an interactive multimedia application for learning the Latin language. In order to evaluate its application, we conducted a research in two sections of second grade levels at the Fifth Grammar School in Zagreb. This paper presents the results of the abovementioned research.

Literature review

A group of Greek scientists (Haung et al., 2011) conducted a research about the usage of multimedia in learning modern Greek language via Internet. The goal was to investigate the effectiveness of this type of learning and to see if this way of learning would make the Greek language more popular among students from Chinese universities. According to the authors, the results indicated that multimedia learning combined with classical tutoring, positively affected the learning environment. Most of the students (95%) were either completely or partially satisfied with the multimedia learning environment. However, 33% of the students declared that human interaction is necessary due to pronunciation difficulties and vocabulary exercises when working from home. Accordingly, we find that the application for learning the Latin language can also be utilized at its best via blended learning, i.e. the combination of traditional and computer assisted learning.

An empirical quantitative research was conducted among college students learning the English language at Qingdao University in China (Liu, 2010). The aim was to investigate the effectiveness of learning the English language with multimedia at Chinese universities. He compared two groups of students: the experimental group which learned with multimedia learning materials only, and the control group which learned with classical teaching methods. Surprisingly, the difference in final test results between the two groups was rather small, showing how the different teaching methods affected the learning outcomes the same way.

A research on the effectiveness of the usage of software for learning English vocabulary was conducted in Croatian elementary schools (Lauc et al., 2007). The research results showed that students were well engaged in the classroom where the multimedia interactive software was used, which was substantiated by better test results, compared with the control group.

In the following chapters we present our research results regarding the development and evaluation of the Latin language learning application.

Research

In order to determine the effectiveness of applying the software in teaching and learning the Latin language, we examined the differences in knowledge acquisition among students of two separate class sections of second grade level in The Fifth Grammar School in Zagreb. Computer-assisted instruction was conducted with the experimental group using the multimedia interactive application for learning Latin, and traditional instruction was conducted with the control group. The validity of the main hypothesis, which claimed that the multimedia application helps students learn the material and evokes more enthusiasm among the students to learn Latin was tested. The research involved the total of 62 second grade students and it was conducted by means of a blinded experiment. None of the students from either of the two groups knew of the research being conducted. However, the students were familiar with the existence of the software, therefore it was not completely new to them. The proficiency level of students in Latin is estimated as equal in both groups according to their former test results and term grades. Both class sections follow the same curriculum and obtain equal test results on average. The study was conducted within a lesson, mainly concentrating on verb forms, where students were supposed to review the content learned so far. The lessons in both groups were substantially equivalent, meaning that the covered content in the experimental group was identical to the content taught in the control group. In this experiment, the independent variable was the introduction of the application for learning Latin in the classroom, and the dependent variable were the final test scores.

Methodology

The multimedia interactive application for learning the Latin language was used by the students in the experimental group. In order to determine the usefulness of the application by using quantitative indicators, a brief knowledge test was conducted at the end of the lesson, with both groups. The students' motivation and interest was tested by means of a survey. At the end of the research the subject teacher was interviewed to give the research review as well as a reflection on the experiment in general.

The software was designed and developed by the authors of this paper within the course *Computer Assisted Teaching* at the Department of Information and Communication Sciences at the Faculty of Humanities and Social Sciences, the University of Zagreb. The application was designed in Adobe Flash, a tool for designing multimedia applications and other components that support user interaction. While designing the application, we were trying to consider a fact that people are counting on computers to supply relevant, truthful, informative, clear, unambiguous and brief information (Blake, 2013). The application consists of four units. The first unit contains the most important grammar rules which are linked with the relevant exercises throughout the software so that students can quickly get help while answering the questions in the second unit.

This way, there is no need to browse books or notebooks for additional reference. In addition to grammar rules, almost every page contains vocabulary hints that describe the words used in the current window. The other two units contain questions about Latin civilization and Latin proverbs. All units are interconnected in order to facilitate good navigation. An important feature of a multimedia software is interactivity which enables every user to choose their own learning path in navigating through information (Skendžić and Kovačić, 2011). Interactivity is highly supported in this application for learning the Latin language.

The main task of this application is to help students master the Latin language curricula as taught in grammar schools in the entire country. The three main sections are the following: 1) exercises that involve many repetitions stretching through three sections of grammar exercises, 2) Latin civilization section and 3) a section with proverb exercises. As the main resources we used the subject teacher’s classroom materials as well as two textbooks (Salopek et al., 1985; Salopek et. al., 1986).

The interactivity in the application is achieved by four types of questions: drag and drop, fill in the blank, true/false and multiple choice. The goal was to motivate the students answer the questions which are formed in an unusual way. The answer checking system provides instant feedback on the correctness of the answer, which is very important to help the learner control the process of learning by either moving forward of referring back. By the simplicity of the design and user friendly navigation, we ensured that students’ attention is not distracted by unnecessary details (as displayed in images 1 through 4).

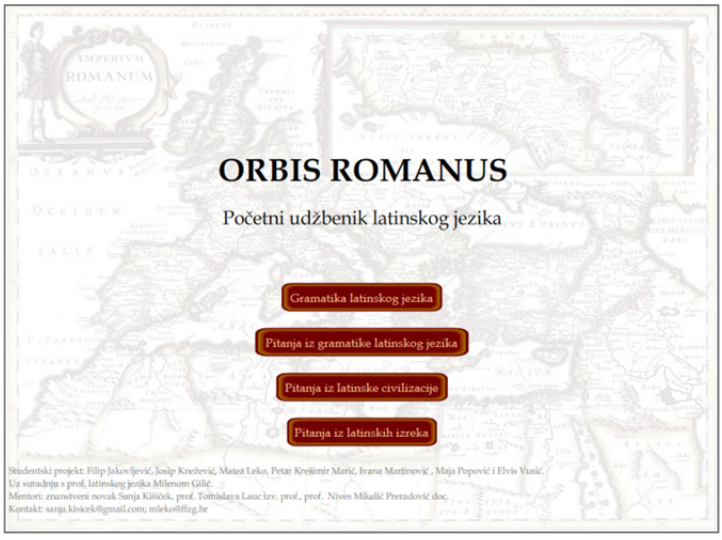


Image 1: Software interface – first frame

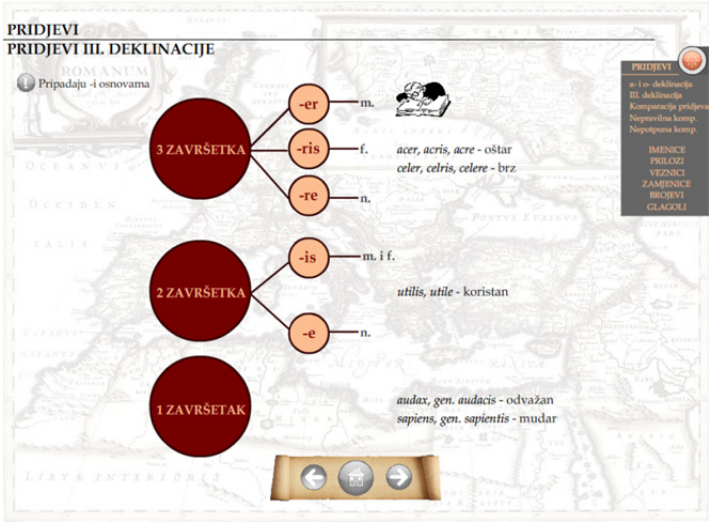


Image 2: Software interface – frame with grammar rules and navigation menu

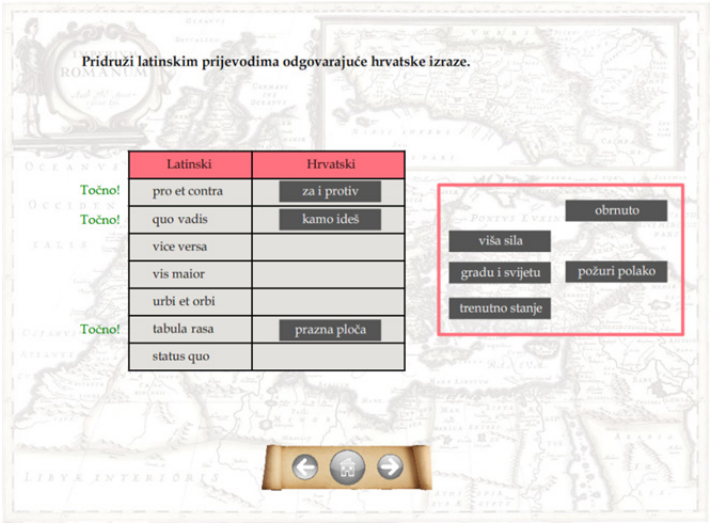


Image 3: Software interface – drag and drop type of question

The hypothesis that students from the experimental group will be more successful in acquiring knowledge than the students from the control group was tested with a short knowledge quiz at the end of lesson with both groups. As well as the learning content was identical in both groups, the knowledge test was also identical for both groups of students, i.e. for both class sections. The students

had to answer questions from the Latin grammar, more precisely verb forms. The maximum score a student could earn by the test was 19 points.

Image 4: Software interface – fill in the blank and multiple choice questions

Motivation and enthusiasm were examined via short questionnaire distributed to students at the end of the experimental lesson after the knowledge test. The survey aimed to confirm additional hypothesis that students from the experimental group will show higher enthusiasm and motivation for learning Latin by using the application than by classical teaching methods. The questionnaire contained four closed types of questions assessed with the Likert scale and one open type of question.

Finally, an interview with the subject teacher was conducted. The aim was to review the entire research and to obtain feedback from the subject teacher as well. The teacher positively reflected on the usefulness of this software, emphasizing the importance of repetition during studying the language. The economy of time very often does not allow extensive repetition to happen in a classic classroom setting with no adequate technology available. Computer software this way fosters individuality in classroom teaching enabling repetition and learning at one’s own pace.

Results

The results of a short knowledge assessment indicate the difference in success between the experimental and the control group. The knowledge test contained

19 questions in verb forms. The results have confirmed the research hypothesis and proved that the experimental group has mastered the learning material more successfully than the control group. The average grade in the experimental group was 3.8 and the 3.3 in the control group which is shown in Table 1.

Table 1: Results obtained by the knowledge assessment

Grade		5	4	3	2	1	Average grade
Group	Number of respondents	Percentage (%)					
Experimental	32	28	44	9	16	3	3.8
Control	30	16	27	40	7	10	3.3

Table 2: Descriptive statistics of achieved scores

Group	MEDIAN	MODE	MIN	MAX	RANGE
Experimental	16	18	7	19	12
Control	14	13	3	19	16

Data in Table 2 shows the difference between the experimental and control group referring the achieved scores. Medium score (median) is 16 in the experimental group, 14 in the control group. Data shows that the most common achieved score (mode) is 18 in the experimental group, and 13 in the control group. Besides that, the minimum amount of scores (7) for the experimental group significantly differs from the minimum score (3) in the control group, while the maximum amount of scores is identical in both groups (19) which means that in the experimental group, as well as in the control group some of the students have achieved a 100%. The span between the minimum and maximum amount of scores is bigger in the control group (16) which results from difference in minimum scores, considering that the maximum amount of scores is equal.

Table 3: Descriptive statistics of achieved grades

Group	MEDIAN	MODE	MIN	MAX	RANGE
Experimental	4	4	1	5	4
Control	3	3	1	5	4

Table 3 shows descriptive statistics of achieved grades. The difference between the experimental and control group is presented by median and mode values. Medium and most common achieved grade in the experimental group is 4, while in control group those values are 3.

Chart 1: GPA in experimental group

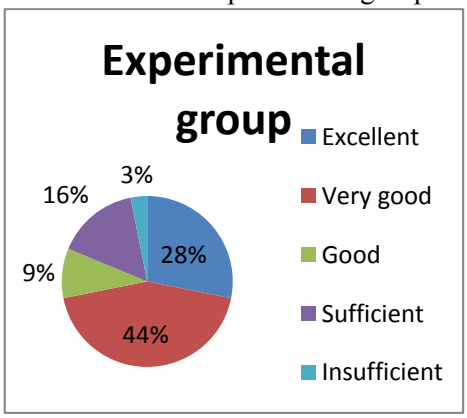
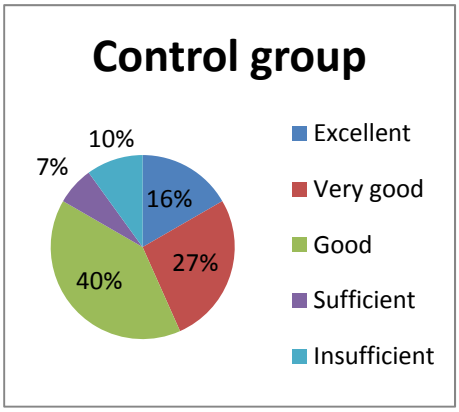


Chart 2: GPA in control group



As displayed in Chart 1 and Chart 2, 72% of students in the experimental group got an excellent or very good grade, while the overall percentage of excellent and very good grades in the control group was only 43%.

Apart from the knowledge assessment, the results of the inquiry in students’ motivation and interest are also significant. The survey on the students’ motivation and interest brought positive results. 88% of the students said they liked using the application for learning Latin, and none of the participants have expressed that they disliked the application (see Chart 3). From these results it is evident that students are open to computer assisted methods of learning.

Similar to the previous results, 84% of respondents agreed or strongly agreed to the statement that learning Latin is more interesting when using the multimedia interactive software. 12% of the respondents have not stated their position while 3% of them disagreed with the statement (see Chart 4).

The third statement aimed to investigate how many students wanted to have such materials for learning Latin for work at home. 72% of students agreed with the statement and would like to have access to these materials at home which is a good incentive for further work. In contrast, 18% of students would not like to use these materials outside of school. It is very likely that the result shows the percentage of students' attitude toward learning Latin in general. However, the fact that not all students have the ability to use computers in their homes should not be neglected.

Furthermore, 66% of the students showed a distinct preference to this type of learning and showed interest for similar tools in other subjects, which can be expressed as the desire to learn, accomplish or stimulate (Vallerand et al., 1992), while 21% have no opinion on the matter.

In the 5th question (which was an open ended question), respondents were given the opportunity to describe their impressions in their own words, give suggestions or criticism. Through the answers to other questions, the majority of respondents expressed positive attitudes regarding the use of the application de-

scribing the application as interesting, useful, great, excellent, creative, fun and easy, noting that learning this way is easier, the material is accepted quicker and well-remembered. In addition, respondents state that learning this way is more creative and useful than by the usual method of learning Latin *ex cathedra*, and several of them have pointed out that it was the first Latin class that was fun. The survey analysis indicates that the students liked the computer assisted lesson in great respect. The survey results showed that the interest and desire for these and similar systems is undoubtedly present.

Chart 3: Students' point of view about the usage of the application

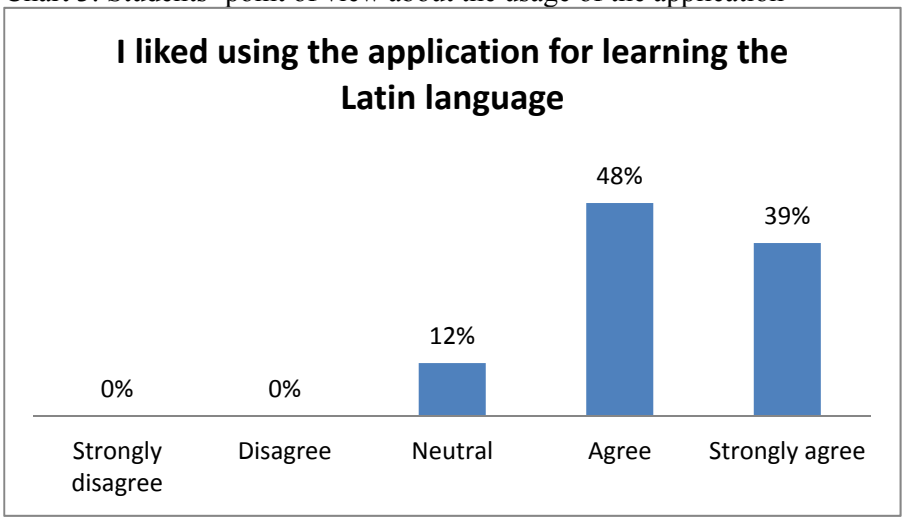
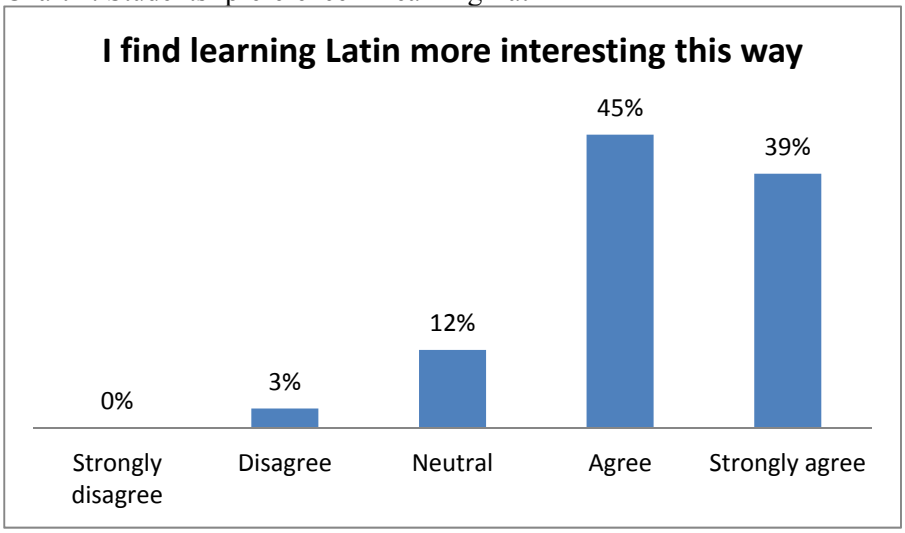


Chart 4: Students' preference in learning Latin



And last but not least, from a short interview with the subject teacher, we learned what it looked like from the teacher's perspective. The teacher confirmed that the results obtained by a short knowledge test were quite expected, both in the experimental group and in the control group, noting that the students in the experimental group were much more interested in the lesson, which was expected to result with better scores. Depending on the needs of the lesson, students of The Fifth Grammar School in Zagreb often use the computer in their Latin class, mostly for presentations and seminars, while they haven't encountered specialized software applications for learning the Latin language so far (although it was only presented to them prior to the experiment). There is a large disparity between the availability of software present for learning languages such as English or German, and those for learning Latin due to their popularity and presence. According to the teacher, there is no doubt that the presence of these tools would increase the popularity of the Latin language among high school students. She especially emphasized the usefulness of the Latin language learning application for grammar repetition and practicing, as part of a lesson which is usually uninteresting and monotonous to students. Although students repeat the learning material through oral and written exams, the teacher also gives great importance to repetition during new lessons. She sees a big potential of using the Latin language learning application. By means of different interactive exercises, repetition is done in an efficient and interesting way. At the end, the teacher claims that without no doubt will this type of learning system become a required part of every teaching in class. She considers that this language learning tool will help the popularization of Latin among students.

Conclusion

The research has shown that the Latin language learning application has helped students with their learning and increased their motivation as well as their enthusiasm for learning the Latin language. Obtained results show that a good relation is accomplished between learning and having fun, the two actions which students frequently experience with different emotions. On one side there is the computer, which they attribute to games, fun and friendships and on the other side there is studying which is mandatory and not always fun. The results point to the importance of motivation in the process of learning, as the students who were using the Latin language learning software have become more interested in the learning material and had better results than students who participated in classical teaching. Motivation is first of all reflected through active participation in lessons which is a postulate for easier language learning.

Due to instant feedback on the correctness of the answers, the possibility to overview the grammar rules and the linkage of different segments within the application, using the Latin language learning application fosters easier and faster learning. Obtained results substantiate the claim that students are prepared

to use applications like this in general, so they give a good rationale for upgrading the application and for further research involving more classroom time over multiple lessons.

Acknowledgements

The authors hereby express appreciation to the Latin language teacher, Milena Gilić, from the Fifth Grammar School in Zagreb, for all the hard work and her involvement in this research. She was working closely with the authors while preparing the content for the multimedia interactive exercises, helped in conducting the experiment, as well as in analyzing results.

References

Blake, R.J., Brave new digital classroom: Technology and foreign language learning. Georgetown University Press, 2013.

Huang, X., Dedegikas, C., Walls, J. *Using Multimedia Technology to Teach Modern Greek Language Online in China*: *Development, Implementation, and Evaluation*. // EURODL, European Journal of Open, Distance and E-learning. URL: http://www.eurodl.org/?article =417 (12th April, 2013)

Lauc, T; Matić, S.; Mikelić Preradović, N. *Project of developing the multimedia software supporting teaching and learning of English vocabulary*. // Proceedings of the 1st International scientific conference "The Future of Information Sciences: INFuture2007 – Digital Information and Heritage" / Seljan, S., Stančić, H. (ed.) Zagreb : Odsjek za informacijske i komunikacijske znanosti, Filozofski fakultet, 2007

Liu, J. *An Experimental Study on the Effectiveness of Multimedia in College English Teaching*. CCSE, English Language Teaching, Vol. 3, No. 1, 2010. URL: http://www.ccsenet.org/journal/index.php/elt/article/view/5255/4353 (1st September, 2013)

Parsons, J. J., Oja, D. *New Perspectives on Computer Concepts*. Introductory. 7th ed. USA , Cengage Learning, 2010. pp. 290 – 292

Salopek, D., Šešelj, Z., Škiljan, D. *Orbis Romanus II*: Latin language textbook. Zagreb, Školska knjiga, 1985

Salopek, D., Šešelj, Z., Škiljan, D. *Orbis Romanus I* : Latin language textbook (beginner level). 4th ed. Zagreb, Školska knjiga, 1986

Skendžić, A., Kovačić, B. *Interactive Whiteboard and Sketchpad Application 4.07 HR – interactivity in the teaching process*. // 4th workshop of the scientific program *Advanced Networks and Services for the Computer Science Society* / Maja Matijašević (ed.), Zagreb, Faculty of Electrical Engineering and Computing, 2011

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., Vallieres, E. F. *The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation* in Education in: Čukušić M., Jadrić, M. E-učenje: koncept i primjena. Zagreb : Školska knjiga, 2012., str. 64.

Media Competence of Students of Maritime Studies in Split

Mila Nadrljanski
Faculty of Maritime Studies – Split
Zrinsko Frankopanska 37, Split, Croatia
milamika60@yahoo.com, mila@pfst.hr

Gorana Tafro
Croatia osiguranje
Trg Hrvatske bratske zajednice 8, Split, Croatia
goranatafro@gmail.com

Ivana Čović
OŠ Manuš
Vukovarska ulica 11, Split, Croatia
covic.ivana@gmail.com

Summary

Media competences comprise a variety of skills related to media (computer literacy, film literacy, digital literacy, etc.), knowledge and skills to critically analyze the media market and the ability to use media in acquiring knowledge. In today's world, where media competences are necessary for communication and collaboration, media education, being the focus of media action, should be given more space in the education of students of Maritime Studies. Since the media have become a powerful industry focused primarily on profit, power and influence, the duty of the competent and of all those involved in the education system is to develop the students' critical awareness in order to protect them from the negative influences of media.

Key words: media competences, media market, critical awareness, Maritime Studies.

Introduction

According to pass rates, the Faculty of Maritime Studies is among the weaker elements of the University of Split. One of the main reasons for the low pass rates is insufficient media competence. The current education system lacks procedures to teach students how to interpret the theory with regard to their area. It is important to know how to use ways of developing arguments, meta-theory and paradigms in empirical research, i.e. use them in contemporary public discourse relevant to the essence of what is being studied.

This paper analyzes the role and the importance of media competence in personal development and success in college. It summarizes the process of defining media competence and the main goal is to determine whether the students of the Faculty of Maritime Studies in Split have mastered the basics of media relations, namely the techniques and tools of media relations.
Research results indicate the way in which students will develop self-awareness, critical self-assessment and willingness to participate and create a creative communication. In this way the students will use special techniques to develop their own personal role in media relations, the ability of placing their own beliefs, communication skills, information management, ability to work in a team and, of course, the ability to apply their knowledge in practice.
Analysis of the survey conducted among a target group of students will determine their attitudes toward individual and mass media in general, interpret personal prejudices, thoughts and experiences with media sources during the course of study.

The notion and the essence of competence

Competence is one of the most controversial terms in the field of general and applied communicology. Competence implies a high level of qualification, i.e. the full integration of knowledge, skills and orientations that are accepted as values of the rules of conduct, which emphasize empathy, respect, teamwork, symmetrical and proactive communication. Contemporary education therefore involves acquiring certain levels of competences, which can be defined as general, specific, specialized, methodological, social and personal.
General competences determine the ability to analyse and synthesize the basic general knowledge in a certain profession.
Specific competences are abilities related to a specific category of skills in the application of knowledge, such as the interpretation of experimental data.
Skills can be defined as specific categories of conceptual, technical, communication, interpersonal, team, intuitive and other abilities, either innate, acquired by experience or achieved with instructive exercise, that improve with consistent application in practice. They are defined with effective methods or ways in which a profession is exercised, and the basis for their effective application is provided by specific aspects of continuous knowledge. Different professions require different skill categories, but the conceptual and communication skills seem to be dominant.
Specialized competences include cognitive abilities to acquire different knowledge content and its integration into a special expertise that belongs to a particular scientific or professional discipline. Specialized knowledge can be applied only if it is based or built on a broad basis of general knowledge.
Social competence can be defined as the ability to establish relationships in communities that operate in the context of a profession, while personal compe-

tences include the ability to mobilize the personality in the exercise of professional activities.

Professional competences in a given area therefore include:

- general, specific, specialized and methodological knowledge of a wide range of action and its effective application, intelligent implementation and wise use,
- a wide range of skills for a wide range of tasks,
- singular detailed skills required for work in a particular area,
- the ability to approach things systematically or to understand the possibilities and limitations from all aspects,
- working in multidisciplinary and multicultural teams,
- the ability to take the initiative in solving problems and in developing new tasks in which a number of issues must be considered.

Methodological competence is the ability to apply specialized knowledge in dealing with professional and scientific issues in research, with the aim of achieving high competence for the successful performance of complex tasks.

The significance of media competence

The word medium indicates different means of communication. A medium implies the materialization of signs and the background of the materialization transmission. The medium gives a specific feature to the code, to its nature, structure and function, so it is often said that the medium is the message. Many scientific publications contain different definitions of the medium. Quoting some of them may help in perceiving the complexity and the importance of this phenomenon in education.

"A medium is a technical and organizational infrastructure in communication. The media include all those instruments and devices that help publishing material intended for the public." (Helbig, 2001)

"The media are the agents of human senses. This implies the place of presentation, the public, as well as transport routes for the transmission of content of a certain capacity." (Schellmann, 2002)

"The medium implies material, mechanical or energy (electrical, electromagnetic, electronic, optoelectronic) links through which data is transmitted. Informational, mechanical and electrical means for data transmission are determined according to three basic phenomena: transmission, processing and storage." (Schellmann, 2002)

The media have an influence on the formation of opinions among people and are very important for the formation of values and attitudes, especially among students. The media evolve cumulatively, i.e. they do not alternate in use – all the media that have been developed throughout history can still be found in use today.

Each medium has its own language. The language of the print media is the written word, the language of the radio is the sound, the language of film is audiovisual language and the language of the Internet is a sophisticated computer communication.

It can be argued that the media are neither harmful nor useful and that they can be both. Although the media is usually seen as a useful source of entertainment and information, the public often deals with dangerous and harmful side of the media (excessive violence, pornography, stereotypes and sensationalism). According to Rotar (2005) both aspects influence social interaction and the formation of people's identities.

The study and the evaluation of the human behaviour led by moral principles can be seen either as the default behaviour that individuals have developed, or as a kind of duty and responsibility that the society requires of its members. This can also be considered through aspects of media competence, especially its meaning.

Modern technology means constant new opportunities in the media industry and constant new requirements for the education of media users. Electronic computers have been transformed from mass data processing devices to devices for information. This is a key moment that changes the very essence of the information system (Nadrljanski, 2012).

Apart from electronic computers, other forms of media – television receivers, digital cameras and camcorders, digital radio receivers, satellite programs, etc. – also require certain knowledge of technology and its use, which in turn requires the users to expand their existing knowledge and skills, i.e. develop media literacy. Interactive television and online radio, cheaper, faster and more diverse versions of their relatives – radio and television – will bring new challenges and adapt to new needs of users, but require new knowledge. Mobile telephony, as an industry on the rise, increases the number of services every day and is becoming a more sophisticated way of human communication. The result is that nowadays there is more and more distance communication and it can be argued that the media have greatly influenced this change.

Changes have occurred in the written media as well. The written word reduces the need of direct interpersonal relationships in the transmission of ideas. One of the disadvantages of text-based communication is that the participants of this type of communication cannot properly assess other people's feelings and moods based solely on text (Nadrljanski, 2010). The Internet and the mobile phone generation faces a language that is no longer the standard language, but short, coded language, with separate sections of text that seem to eliminate all previous ideas of text-based communication. What was formerly recognized as the main body of text has been transformed into independent message structures.

Since it is only the text that defines reading and readers, and since text is no longer read but the message, it is possible to talk about recipients rather than

readers. Unusual symbols, abbreviations and funny signs are intended to provide full information and shorten time as much as possible. It can be argued that this is the absurdity of today's society – people now have everything within their reach, but lack the time to explore and to manage it all.

The technological aspect of media literacy is not the only one, as informatization and computerization are only parts of media literacy. Apart from learning about all media, their history, production and economic principles of functioning, the modern concept of media literacy also includes some important issues such as the question of who owns and controls the media and what are the consequences of media ownership (the concentration of social power, the filtering and the commercialization of content, etc.).

The media have a great social responsibility when it comes to acting in the public interest and the needs of individuals, promoting diversity and quality of program. However, the question of whether the media act accordingly is somewhat questionable, if not controversial, and it is for that reason that the modern concept of media competence must contain principles related to critical assessment.

Achieving media competence

The importance of media education is found in the definition proposed by the International Council for Film and Television in 1973 (Nadrljanski, 2010): "Media education should include studying, teaching and educating with the help of modern means of communication, which are becoming an integral part of a specific and autonomous area of the cognition process in educational theory and practice. Attention should be paid to the different ways of using them as auxiliaries in teaching and education, but also in other areas of cognition."

"Media education" namely involves the acquisition of skills for critical consideration of any kind of media (newspapers, radio, television, etc.). The aim is to reduce the distance between the media and the users, to understand how the media function and to become familiar with their contents, as well as their way of placing themselves in different perspectives in relation to systems in which they are being developed.

The term "media education" was introduced in the 1960s in international circles dealing with problems of education, especially in circles related to UNESCO. In this period the predictions of experts focused on the apparent explosion of mass communication, especially on television. Many topics were discussed at the time, the most common of which will be mentioned here: the power of the new "magical" device in teaching literacy to the masses not included in the existing educational structures, the reluctance of teachers to embrace television as a legitimate approach to knowledge and the need for a critical assessment of hazards of media manipulation.

In all these cases, it seemed necessary to gain specific knowledge that would enable an impartial consideration of the problem. "Media education" thus gave the ability to respond to multiple expectations. However, from a range of differ-

ent issues a dominant one quickly emerged – the issue related to the method of accessing the reality through the media in order to gain knowledge and put it to good use, but in such a way that the student is taught how to distance himself enough from that reality.

Len Masterman (1985) lists seven basic reasons for the necessity of media education:

- high consumption of media and the satiety it has caused
- ideological character of the media, especially advertising
- emergence of information management in enterprises (government offices, political parties, ministries, etc.)
- increased infiltration of the media in democratic processes (elections are primarily media events)
- increased importance of visual communication and information in all areas of human life and work (except schools, which give priority to the printed material; communication systems are mainly visual)
- expectations of young people to be educated so that they can understand the period they live in (what is the meaning of education and the growth of individual culture if technological means are so carefully avoided, as is the questioning of values in today's life?)
- national and international growth of privatization of all information technologies (when information becomes product, its role and its properties are changed).

The motivation of students can be influenced from the outside and the inside. The exterior aspect of motivation is largely determined by the society, based on the known factors of motivation, such as employment, social and financial status, the possibility of promotion, and others. Interior motivation and the motivation during the studies will be greatly managed by curriculum, the quality of the educational process, working conditions and equipment (the media) and the conditions for a creative approach to work.

Teaching and learning processes are influenced by many components, but basically they mostly depend on individual abilities of students. Objectively, it is possible to improve the processes with selection of program contents, new methods of teaching and learning and the introduction and use of modern media technology.

Until recently university education has been considered as a level towards which students should not be encouraged. Changes in the European education system show that special attention is given to motivation. As one of the basic educational methods, encouragement has so far been reserved for primary and secondary schools, but today there are a number of researches and theorists who suggest that encouragement can and should be effectively used on a university level.

Methods of encouraging are being developed and its means are being adapted to the needs of students. The means can be divided into four areas: cognition, emotion, socialization and the active area. It is necessary to elaborate a number of means of encouragement in each area and university professors should be familiar with all of them.

The use of information and communication technology may well become the latest teaching aid in all areas of education. The possibilities of multimedia presentations and access to computer networks, especially the Internet, provide immediate access to a vast and rapidly growing number of information from all over the world. The importance of educating students for the quality use of information from the Internet has the purpose of acquiring knowledge. Notions of critical literacy and information literacy are essential for understanding the functioning of an individual in the information space available over the Internet and the skills of critical thinking are the basis for critical and information literacy.

Dynamic changes in the contemporary society have necessarily stimulated the need to change the existing fundamental tasks of universities and other institutions of higher education. The goal is to find a new paradigm that complies with the fundamental characteristics and requirements of the information society and also to find a new paradigm of a contemporary university library.

Enhanced by rapid technological and educational development of the education system, changes in higher education are the result of global processes such as:

- the democratization of education and a significant increase in the number of students
- interdisciplinary approach
- increasing diversity in age and in academic profiles of students
- the orientation of society towards lifelong learning
- collaboration between contemporary universities and the industry, where knowledge and information resources are the driving forces of companies
- more emphasis on research by the academic staff and the development of the university as a centre of scientific research, especially in smaller communities
- growth of information and communication technologies that enable a number of updated information to those who have the skills to access them and to use them.

Faced with these challenges, the universities respond by re-evaluating their main goals and visions as institutions of higher education and their previous viewpoints on teaching and education, where the focus of attention is on learning and students. The tendency among universities is to convert from the traditionally structured organizations to so-called open universities, e-universities or consortia of higher education institutions, private sector and other institutions.

This emerging type of university is under the influence of the concept of virtual learning and uses the Resource Based Learning approach to education.
With the birth of the concept of the hybrid library which is based on the awareness of the importance of immediate access to information and information sources, not only in the library but also in classrooms, the entire campus and users' homes, university libraries and other information services are transformed into an active factor in the process of teaching, education and research. Libraries assume a fundamental role in the process of distance learning, particularly in relation to the construction and maintenance of electronic archives. University libraries will play a key role in supporting and improving access to network resources.
Although electronic sources will become more important for the archives of these libraries, especially those related to distance learning obtained from professional companies or created at the university, the management, development and protection of printed material will not lose its importance. Working within conditions of constant change, university libraries should accept excellence as the foundation of its development by building their own systems of improving quality and keeping the orientation to the end user with a constant individualization and differentiation of services.

Research results

A survey on media was conducted with students of the Faculty of Maritime Studies in Split who attended the course Communicology. The survey was conducted on a sample of 47 students aged 21 to 22. This is how the surveyed students relate to media:
Students read magazines once a week on average, mostly for entertainment. Among daily newspapers "24 hours" is read the most, by 87%. Even in the daily news entertainment is the most interesting section; most students do not read politics. They do not buy newspapers and magazines every day. Only 5% of the surveyed students read professional journals on a regular basis, 64% read professional journals occasionally and 31% do not read professional journals at all.
The respondents watch television three hours a day on average. Entertainment is the most preferred television content (93%), while politics is the least preferred (4%). 38% of the surveyed students use Internet on a daily basis, while 14% do not use Internet at all.
Interestingly, 63% of students believe that the media are unethical in most cases and 42% believe that political interest is related to the issue of ethical/unethical behaviour in media.
When asked whether they think the media affect their cognitive abilities, i.e. whether they believe that the media lessen intellectual abilities, the power of critical thinking and emotional engagement, 35% replied yes, 28% no, while 37% said that they are not sure. Additionally, 78% of the respondents believe that there is violence in media and over half of those respondents think that the

violence is manifested through deliberate misleading of the public by displaying modified and subsequently processed information.

Conclusion

The importance of developing critical thinking skills in the area of critical reception and in the area of critical production should be pointed out. Emphasis should be put on the role that computer science teaching plays in supporting the development of critical thinking for quality use of information from the Internet. It can be concluded that full effects can be expected only after careful correlation of computer science and communicology with other courses at the Faculty of Maritime Studies in Split. However, the application of critical thinking skills to use information from the Internet is expected regardless of faculty courses.

References

Helbig, Elke; Köhler, Winfried; Lümkemann, Andre. Medienpädagogik. Neusäß: Kieser, 2001

Masterman, Len. Teaching the media. London: Comedia, 1985

Nadrljanski, Đorđe; Nadrljanski, Mila. Informacijsko-komunikacijski sistemi i tehnologije = Information and communication systems and technologies. Split: Redak, 2012.

Nadrljanski, Mila. Komunikologija i menadžment: istraživanja komunikoloških aspekata obrazovanja menadžera = Communicology and Management: Research on communicational aspects in the education of managers. Split: Redak, 2010.

Schellmann, Bernhard. Medien verstehen, gestalten, produzieren. Haan-Gruiten: Auflage, 2002

Zgrabljić Rotar, Nada. Mediji - medijska pismenost, medijski sadržaji i medijski utjecaji = The media - media literacy, media content and media influences. *// Medijska pismenost i civilno društvo = Media Literacy and Civil Society*. Sarajevo: Mediacentar, 2005.

Computer-based Assistive Technologies in Education for Students with Disabilities

Marko Arambašić
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
marambas@ffzg.hr

Ivan Dunđer
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
ivandunder@gmail.com

Summary

The aim of this paper is to accentuate the role of computer-based assistive technologies in the education process of students with disabilities. The paper presents preferred assistive technologies and types of educational problems that disabled students face. Impact of assistive technologies on independence in education and on quality of studying was also analyzed. Since the focus of this paper lies on the usability of computer-based assistive technologies, as fundamental elements of the education of disabled students, the importance of their accessibility is being particularly emphasized.

Key words: computer-based assistive technology, education, students with disabilities, speech technology, communication, input technology, Barthel index

Introduction

Today, people use technology to function more completely and efficiently in their lives. However, for people with physical disabilities, it is very often impossible to function in a world designed for people without disabilities, so they take advantage of a variety of methods to gain access to information technology and computer-based assistive technologies for (wireless) communication, mobility and daily living tasks. According to Raskind (2000), assistive technology can be defined as any item, piece of equipment or system that helps people bypass, work around or compensate for learning difficulties, which cannot be cured or outgrown. Learning disabilities are professionally diagnosed learning difficulties with reading, writing, speaking, listening, spelling, reasoning or math that are the result of a presumed central nervous system dysfunction. Nowadays, the computers are needful instruments for students with disabilities,

offering them a new perspective and a new way to live and learn. Many people with severe physical disability but normal speech (e.g. spinal injuries) use speech recognition as a mean to input text, as it can give faster input rates than adapted keyboards (Hawley et al., 2005). Because disabilities differ among students, each student must be fitted with assistive technologies that are commensurate with their individual needs (Christmann and Christmann, 2003). For such users, the products which have existed on the market offer additional accessibility to computers and are created for each type of disability (Isaila and Smeureanu, 2008). For example, assistive technology for visually impaired students includes screen enlargers and text readers, speech recognition systems, text-to-speech synthesizer, assistive input technologies, such as optical character recognition (OCR), word prediction programs, etc. Assistive products for persons with mobility disabilities include speech technology systems, wireless communication technologies, editing programs on screen using alternative electronic products, alternative keyboards, keyboard filters for editing or touch screens. Students with learning impairments can use word prediction and reading comprehension programs or speech technologies.

Speech technologies

Speech technology is potentially of enormous benefit to students with unintelligible speech (or no speech) and therefore has an important role in support for spoken communication. In order to be successful, speech technologies should effectively take into account the needs of user groups and have the ability to adapt to the needs of individuals. Communication aids range from simple devices which play back a small number of messages stored as recorded speech, to very sophisticated devices allowing access to large annotated vocabularies with synthesized speech output (Hawley et al., 2005). Recent speech technology research for Croatian language has concentrated upon speech synthesis (Boras and Lazić, 2006) and particularly on domain-specific evaluation of synthesized speech in order to increase the naturalness of artificial speech (Dunđer et al., 2013).

Text-to-speech synthesis

Text-to-speech systems have a vast range of applications. First real use was in reading systems for visually impaired people, where a system would read some text and convert it into robotic-sounding speech. Today, sophisticated systems can help impaired users to navigate around a computer system (Taylor, 2009). Text-to-speech synthesis can also be applied in an automated dialogue system, which then allows users to conduct entire financial or similar transactions. Reading systems for reading news stories, weather reports and travel directions are also based on this technology. They convert text that is displayed on the computer monitor into speech, allowing students to gain independent access to assignments, books, and learning material. Teachers or students do, however,

need to pre-scan material before they can use it (Wade-Woolley, 2005). Such systems have also shown to be useful to individuals with cognitive disorders and communicative impairments, especially for practicing writing and reading (MacArthur et al., 2001). In addition, research has shown that the use of this technology can actually improve word-recognition and decoding skills (Torgesen and Barker, 1995). Specific synthesis tools, e.g. screen readers are used to transform a graphic user interface (GUI) into an audio interface by verbalizing and converting every object on the computer screen including text, graphics, control buttons and menus into a synthetic voice that is spoken aloud.

Speech recognition

Speech recognition is a complex process, in which a sound is converted into electric signal, processed and then transformed into text. It can help physically disabled students to control a computer via voice trough a microphone or to input text. It is also useful for quickly writing down ideas (De La Paz, 1999), for practicing writing, spelling, reading comprehension and word-recognition (Higgins and Raskind, 2000). The speed of speech recognition also gives it a potential advantage over other input methods commonly employed by physically disabled students. But, it is also the case that many students prefer the non-speech alternative as they find speech recognition frustrating, due to less than perfect recognition or a nonintuitive way of composing text (Hawley et al., 2005). Speech recognition software works with most word processing systems, but a user has to train the computer to recognize voice patterns and pronunciations by reading specific text. The more a user uses a speech recognition system, the better it gets, eventually reaching sufficient accuracy. It can be particularly helpful to individuals whose oral language skills exceed their written production. Although, speech recognition is most useful for students who are verbally fluent, with daily and supervised use it can also have a positive influence on the performance of less verbally fluent students (Wetzel, 1996).

Wireless communication technologies

Use of computers for communication and networking activities via the internet can expand the learning environment beyond the walls of the classroom and allows students with disabilities, just like other students, to conveniently access and send information anytime and anywhere, without constraints of time or place (Hasselbring and Williams Glaser, 2000). Communication technologies become a valuable tool for learning if they offer disabled students opportunities to gather a wide variety of resources and information. Networked through the internet, such a collaborative learning environment enables students to practice communication skills without fear of being stigmatized because of their disability and to exchange ideas, information and knowledge with others. Mobile devices can complement and add value to the existing learning models (Motiwalla, 2007). However, mobile technology is limited by screen size, computa-

tional power, battery capacity, input interface and network bandwidth (El-Hussein and Cronjé, 2010). Nevertheless, e.g. Bluetooth enables wireless data transmission and offers a connection with two or more devices over a short distance. This makes it ideal for a portable and free of charge learning environment, which allows students to actively participate in the class and answer instructor’s questions in an easier way.

Assistive input technologies

Assistive input technologies are designed to provide additional computer accessibility to individuals who have physical or cognitive difficulties, impairments or disabilities and allow individuals to control their computers through means other than a standard keyboard or a pointing device, using alternative (on-screen) keyboards, special pointing devices, sip-and-puff switches, wands and sticks, joysticks, trackballs, touch screens, eye or head movement and eye gaze systems, light-sensitive or pressure-sensitive systems, speech/voice-activated systems, word prediction tools etc. (Obiozor, 2010).

Optical character recognition (OCR)

Optical character recognition (OCR) systems, when combined with speech synthesis, might be used as reading machines. The OCR enables users to input hard copy text directly into a computer. Then the speech synthesizer reads the text back out loud. In this way, user can hear as well as see the text (Raskind, 2000). OCR combined with speech synthesis can be particularly helpful to students who have relatively few problems comprehending spoken language, but have great difficulty with decoding of text (Montali and Lewandowski, 1996). The OCR works with a scanner or similar devices, e.g. reading pens, which are primarily designed to read aloud single words, rather than full sentences (Higgins and Raskind, 2005). The scanner reads printed material, converts it to a computer file and then shows it on a computer screen. OCR software recognizes text by analyzing the structure of the object that needs to be digitized, by dividing it into structural elements and by distinguishing characters through comparison with a set of pattern images stored in a database and built-in dictionaries. This allows conversion of scanned input text from bitmap format to encoded text. However, errors are unavoidable in optical character recognition, and the noise induced by these errors presents a serious challenge to later-stage usage of data (Seljan et al., 2013). OCR systems are available as self-contained units or as systems which work together with computers.

Touch screen technology

Touch-sensitive screens are popular with young computer users and with individuals who have severe developmental or physical disabilities (Hasselbring and Williams Glaser, 2000). This technology allows users to simply touch the computer screen to perform a variety of tasks.

Word prediction

Word prediction programs work together with word processors. These programs predict the word a person wants to enter into the computer (Mirenda and Turoldo, 2006). The person types the first letter of a word, and the program, e.g. T-9 (Text on 9 keys) predictive text technology offers a list of words beginning with that letter. If the desired word appears, it can be chosen from the list by pressing the number on the keyboard that is displayed next to that word or by pointing and clicking with the mouse. That word will automatically insert into the sentence. If the desired word does not appear on the list, the user continues to type the next letter until it does appear. After the user chooses a word, the computer predicts the next word in the sentence. Again, it offers a list of possible words, even before the first letter is typed. Predictions are based upon the sentence content and spelling, as well as the number of times a word is used (Raskind, 2000). Word prediction may be helpful to individuals who have problems with keyboarding, spelling or grammar. These programs may also assist people who struggle to come up with the exact word they want to use in a sentence (Tumlin and Wolff Heller, 2004).

Research and methods

The main hypothesis of this paper is that students with disabilities at the University of Zagreb use computer-based assistive technologies for educational purposes on a daily basis. The idea of this research was to obtain a general overview of the students' satisfaction with assistive technologies, but also to identify characteristics and problems in the process of education of this target group. The research was carried out among undergraduate, graduate and postgraduate students with disabilities at the University of Zagreb. Conducting an online survey was chosen as the research method. The survey method comprises completing an anonymous online questionnaire with 23 open-ended and closed-ended questions. The research was performed in June 2013 on a sample of 10 students; therefore this research is based on ten complete questionnaires. Representativeness of the research sample might not be ensured due to low sample size, but still, representativeness of a sample is not guaranteed by its size. Furthermore, there are is no official and up-to-date data on the number and profiles of students with disabilities at the University of Zagreb. The questionnaires were analyzed quantitatively and qualitatively via content analysis, in order to decipher key problems mentioned among the participants. They were asked to undertake a questionnaire concerning how they consume and perceive computer-based assistive technologies for educational purposes and especially focusing on earlier experience, usability of computer-based assistive technologies, and what technology they find most useful regarding their type of disability. Respondents were also asked to describe their type of disability using Barthel index, which assesses self-care, mobility and continence. Barthel index allows the standardization of treatment and follow-up records, and these scores have

been reported to have a direct relation with the degree of independence of a patient (Wolfe et al., 1991).
In this research authors used a modified Barthel index scale ranging from:

- severe disability - constant nursing care, attention, bedridden, incontinent,
- moderately severe disability - unable to attend to own bodily needs without assistance, and unable to walk unassisted,
- moderate disability - requires some help, but able to walk unassisted,
- slight disability - able to look after own affairs without assistance, but unable to carry out all previous activities,
- no significant disability - able to carry out all usual activities, despite some symptoms.

Results and discussion

Table 1 shows research statistics and first results of the conducted online survey. Most respondents were male (90%) and were coming from different fields of study - Economics, History and Rehabilitation science were mostly represented. Average age was cca. 29 years and 50% of them were still undergraduate students. Among the respondents were also three postgraduate students. 70% of all respondents had moderately severe disability, 2 respondents had slight disability and only 1 no significant disability. All of the respondents had prior experience with computer-based assistive technology, while most of them use free technologies (40%) on a daily basis. 30% uses a combination of free and commercial tools, while one respondent uses only traditional non-electronic technologies. Why this respondent uses only traditional technologies is unclear, but this opens a new question about accessibility of computer-based assistive technologies for disabled students in the education process (price, effort, usability problems, additional training).
Chart 1 shows that even 70% of all respondents had prior experience with speech technologies, mostly using text-to-speech systems in form of screen readers, and speech recognition. Half of respondents had also prior experience with wireless communication technologies, e.g. Bluetooth, which they mostly used for transferring data, homework assignments, for sharing and retrieving study-related information and for taking notes virtually on different electronic devices. 50% of the respondents have also experienced working with assistive input technologies, mostly OCR for digitizing learning material, on-screen keyboards, but also word prediction tools for speeding up text input.

Table 1: Research statistics and results

Respondents	**male**	**9**	
	female	1	
	total	**10**	
Average age of respondents		28,6 years	
Education level		Undergraduate	**5**
		Graduate	2
		Postgraduate	3
Field of study		Economics	**2**
		History	**2**
		Rehabilitation science	**2**
		Accounting and audit	1
		Information and communication sciences	1
		Public administration and public finances	1
		Sociology	1
Barthel index		severe disability	0
		moderately severe disability	**7**
		moderate disability	0
		slight disability	2
		no significant disability	1
Prior experience with computer-based assistive technology		yes	**10**
		no	0
Type of computer-based assistive technology you use		free	**4**
		free and commercial	3
		commercial	2
		none (using traditional technologies)	1

Source: Research conducted in June 2013.

Chart 1: Prior experience with specific computer-based assistive technologies

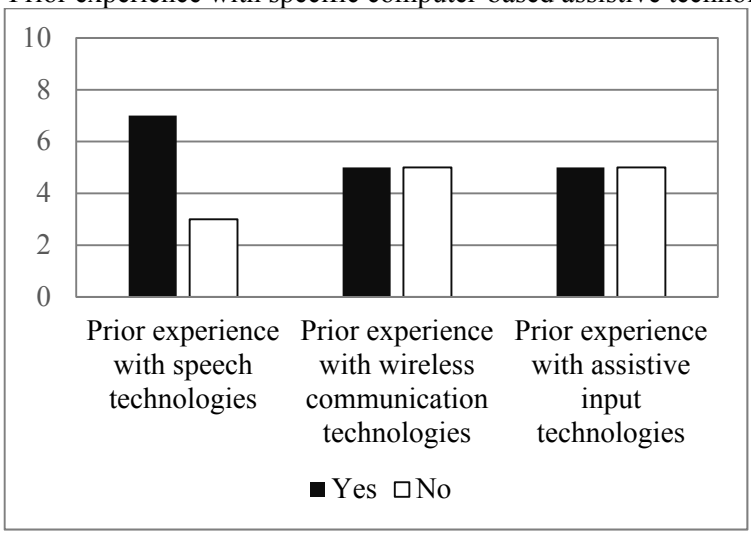
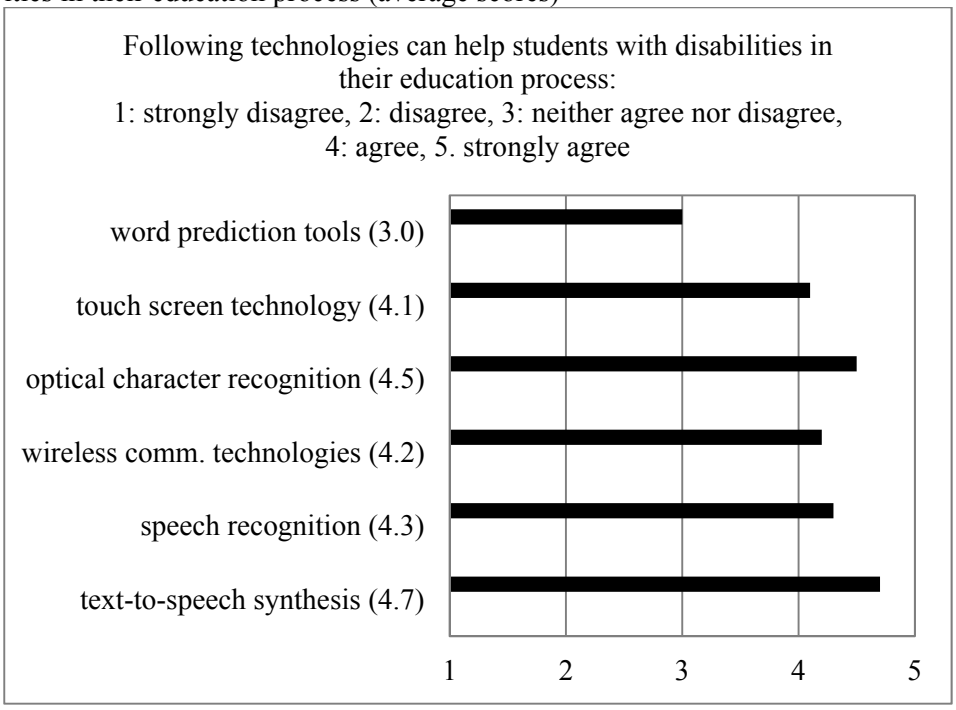


Chart 2 shows how students with disabilities at the University of Zagreb perceive different types of computer-based assistive technologies. The chart is based on Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree) and shows that text-to-speech synthesis (4.7) and optical character recognition (4.5) scored best. In other words, according to our target group those two technologies have more educational advantages than other mentioned technologies. Word prediction tools scored worst (3.0), indicating that students neither agree nor disagree with the claim that word prediction tools can help in the education process.

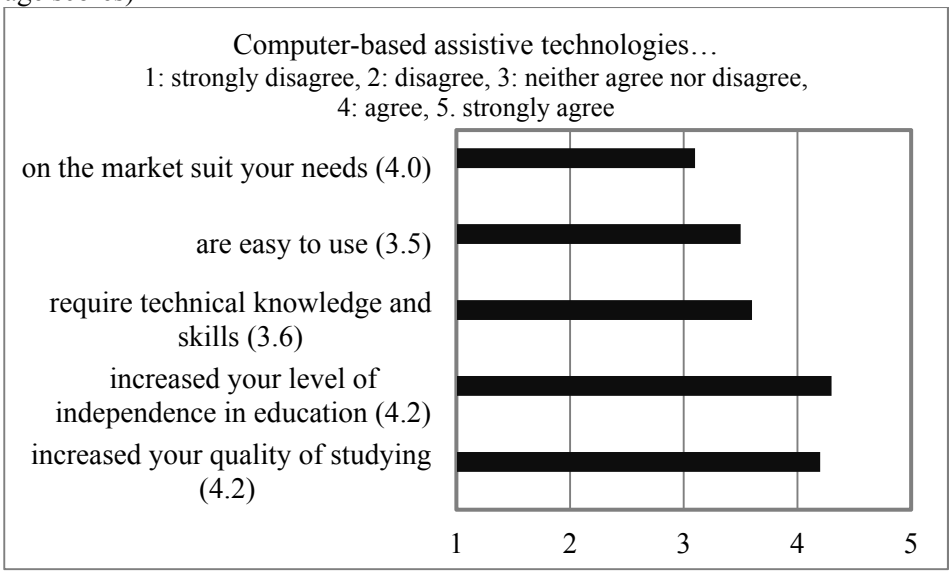
Chart 2: Useful computer-based assistive technologies for students with disabilities in their education process (average scores)



Furthermore, respondents were mostly satisfied with the assistive technologies on the market and their usability, but pointed out that for the effective use, technical knowledge and specific skills are needed, as shown in chart 3. They also claimed having increased their level of independence in education (4.2) and quality of studying (4.2) by using computer-based assistive technologies.
Eventually, respondents were asked to name a computer-based assistive technology that they found the most useful during their education process. 60% answered optical character recognition (3 respondents) and wireless communication technologies (3 respondents), followed by speech technologies (40%) –

text-to-speech synthesis (2 respondents) and speech recognition (2 respondents). Touch screen and word prediction technologies were not mentioned.

Chart 3: Computer-based assistive technologies currently on the market (average scores)



Conclusion

Assistive technology is a type of technology that is available for people with disabilities. Basically, anything that makes a task easier to perform is considered assistive technology, while computer-based assistive technologies are supported by a computer, or a similar electronic device. In this paper, a research in form of an online survey was conducted on a sample of ten students with disabilities at the University of Zagreb. Seven students out of ten had moderately severe disability according to Barthel index. The research gave valuable information on type of assistive technologies used by students in order to enable and enhance education. Furthermore, it was shown that all respondents had prior experiences with computer-based assistive technologies and 90% of them used those technologies on a daily basis. Therefore, the main hypothesis was confirmed. The authors showed relevant information on type of preferred assistive technologies in education: speech technologies and OCR, followed by wireless communication and touch screen technologies. Also most of the respondents favored non-commercial technologies and stated that computer-based technologies increased their level of independence in education and quality of studying. This underlines the huge importance of accessibility of computer-based assistive technologies for disabled students in their education process. Generally, a very positive attitude towards assistive technologies was

noticed, i.e. most of the scores were between 4 (agree) and 5 (strongly agree). Further research on a larger target group is planned in order to uncover real problems in the usage of existing computer-based assistive technologies and to identify space for improvements or modifications.

References

Boras, Damir; Lazić, Nikolaj. Aspects of a Theory and the Present State of Speech Synthesis. // *29th International Convention MIPRO: CTS - Computers in Technical Systems* / Budin, L.; Ribarić, S. (ed.). Opatija : MIPRO, 2006, pp. 187-190

Christmann, Edwin P.; Christmann, Roxanne R. Technologies for Special Needs Students. // *Science Scope*, 26 (2003), 6; pp. 50-53

De La Paz, Susan. Composing via dictation and speech recognition systems: Compensatory technology for students with learning disabilities. // *Learning Disability Quarterly*. 22 (1999), 3; pp. 173-182

Dunđer, Ivan; Seljan, Sanja; Arambašić, Marko. Domain-specific Evaluation of Croatian Speech Synthesis in CALL. // *Recent Advances in Information Science 2013 proceedings of the 7th European Computing Conference.* Dubrovnik : WSEAS Press, 2013, pp. 142-147

El-Hussein, Mohamed Osman M.; Cronjé, Johannes. C. Defining Mobile Learning in the Higher Education Landscape. // *Educational Technology & Society*. 13 (2010), 3; pp. 12-21

Hasselbring, Ted S.; Williams Glaser, Candyce H. Use of Computer Technology to Help Students with Special Needs. // *The Future of Children*, 10 (2000), 2; pp. 102-122

Hawley, Mark S.; Green, Phil; Enderby, Pam; Cunningham, Stuart; Moore Roger, K. Moore. Speech Technology for e-Inclusion of People with Physical Disabilities and Disordered Speech. // *INTERSPEECH*. Lisabon : ISCA, 2005, pp. 445-448

Higgins, Eleanor L.; Raskind, Marshall H. Speaking to read: The effects of continuous vs. discrete speech recognition systems on the reading and spelling of children with learning disabilities. // *Journal of Special Education Technology*. 15 (2000), 1; pp. 19-30

Higgins, Eleanor L.; Raskind, Marshall H. The Compensatory Effectiveness of the Quicktionary Reading Pen II on the Reading Comprehension of Students with Learning Disabilities. // *Journal of Special Education Technology*. 20 (2005), 1; pp. 31-40

Isaila, Narcisa; Smeureanu, Ion. Assistive Technologies. // *Informatica Economica*. 8 (2008), 2; pp. 135-139

MacArthur, Charles A.; Ferretti, Ralph P.; Okolo, Cynthia M.; Cavalier, Albert R. Technology applications for students with literacy problems. // *The Elementary School Journal*. 101 (2001), 3; pp. 273–299

Mirenda, Pat; Turdolo, Kristen. The Impact of Word Prediction Software on the Written Output of Students With Physical Disabilities. // *Journal of Special Education Technology*. 21 (2006), 3; pp. 5-12

Montali, Julie.; Lewandowski, Lawrence. Bimodal reading: Benefits of a talking computer for average and less skilled readers. // *Journal of Learning Disabilities*. 29 (1996), 3; pp. 271-279

Motiwalla, Luvai F. Mobile learning: A framework and evaluation. // *Computers & Education*. 49 (2007), 3; pp. 581-596

Obiozor, Williams Emeka. Technology, Quality Learning and Student Disabilities: Challenges for Nigerian Teachers in the 21st Century. // *International Journal of Educational Research*. 11 (2010), 4; pp. 6-26

Raskind, Marshall. Assistive Technology for Children with Learning Difficulties. San Mateo : Schwab Foundation for Learning, 2000

Seljan, Sanja; Dunđer, Ivan; Gašpar, Angelina. From Digitisation Process to Terminological Digital Resources. // *36th International Convention MIPRO: CIS - Intelligent systems* / Biljanović P. (ed.). Opatija : MIPRO, 2013, pp. 1329-1334

Sharma, Reena F.; Wasson, Geetanjali S. Speech Recognition and Synthesis Tool Assistive Technology for Physically Disabled Persons // *International Journal of Computer Science and Telecommunications*. 3 (2012), 4; pp. 86-91

Taylor, Paul. Text-to-Speech Synthesis. Cambridge : Cambridge University Press, 2000

Torgesen, Joseph K.; Barker, Theodore A. Computers as aids in the prevention and remediation of reading disabilities. // *Learning Disability Quarterly*. 18 (1995), 2; pp. 76-87

Tumlin, Jennifer; Wolff Heller, Kathryn. Using Word Prediction Software to Increase Typing Fluency With Students With Physical Disabilities. // *Journal of Special Education Technology*. 19 (2004), 3; pp. 5-13

Wade-Woolley, L. Education for All. The Report of the Expert Panel on Literacy and Numeracy Instruction for Students With Special Education Needs, Kindergarten to Grade 6. Ontario Ministry of Education 2005

Wetzel, Keith. Speech-recognizing computers: A written-communication tool for students with learning disabilities? // *Journal of Learning Disabilities*. 29 (1996). 4; pp. 371-380

Wolfe, Charles D.; Taub, Nick A.; Woodrow, E. J.; Burney, P. G. Assessment of scales of disability and handicap for stroke patients. // *Journal of the American Heart Association*. 22 (1991), 10; pp. 1242-1244

Students for Seniors: Basic ICT Education for the Elderly

Marija Lenić, Anita Kustura, Ida Jurković, Nives Mikelić Preradović
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
mlenic@ffzg.hr, akustura@ffzg.hr, ijurkovi3@ffzg.hr, nmikelic@ffzg.hr

Summary

The mission of this project was to teach the elderly citizens the basic information and communication technology (ICT) skills. Students of information and communication sciences (museum studies and heritage management and librarianship) planned and realized this service learning project through interactive collaboration between Tin Ujević public library, Trešnjevka retirement home and Faculty of Humanities and Social Sciences in Zagreb. The education for the elderly, a social group that is least connected with the rest of the today's society through information technology, represents a very important issue for their social inclusion and staying in life's mainstream. The students taught the end beneficiaries the basics of information technology with the emphasis on social networks (Facebook and Skype). Out of 15 participants, the first level (basic computer skills) was successfully completed by 13 participants. Two participants had physical deficiencies (they could not coordinate the movements of their body with the visual input). The second level was completed by 9 participants, while 6 participants were able to complete the third level (creating and managing the Facebook account). The last level (creating and managing the Skype account) was successfully completed by only two participants.

Key words: education, Facebook, Skype, information and communication technology, the elderly

Introduction

In the today's society, the Internet is still unavailable to a large part of the elderly population, even though the research results show that information and communication technology (ICT) considerably improves their everyday activities and independence and supports their social inclusion[1].

The elderly are often marginalized in terms of their social influence, since they are less economically stable than the younger population and their health condi-

[1] Lewis S, Ariyachandra T. Seniors and online Social network use. CONISAR 2010: 1-15.

tions are decreased. The situation of women, who usually outlive the men, is especially difficult, since most of them are in a worse socio-economic position[2]. Depression is another condition common in the elderly population due to the loss of spouse, friends and their physical abilities[3].

Society often forms stereotypes about the elderly, focusing on the negative characteristics of the third-age individuals, such as physical and intellectual disability. As a consequence, such a prejudice enables a certain type of social exclusion.

The basic goal of the project *"Students for seniors: Basic ICT education for the elderly"* was to teach the elderly the basics of information and communication technology in a comprehensible way. The project was designed and implemented by four graduate students of Information and Communication Sciences study at the Faculty of Humanities and Social Sciences (who are also authors of this paper) as part of the course *Service Learning in Information Sciences*.

The students in this project aimed to analyze if ICT education can enrich the lives of elderly and positively influence their social inclusion by expanding their personal social networks. (Social inclusion is defined as a condition where people feel valued, their differences are respected, and their basic needs are met so they can live in dignity)[4].

There are various aspects of ICT education for elderly, such as browsing the Internet, sending e-mails, reading news, sending postcards, taking care of personal banking, etc.[5] Students, who planned and realized the project, taught the elderly the basic information retrieval skills on the Internet, how to use e- mail and Skype and the basics of communication on social networks (more precisely, Facebook), since research results show that social networks can positively affect elderly people's well-being, because they strengthen their mental and physical health, they feel safer and more secure[6].

---

[2] Laklija, M.; Rusac, S.; Žganec, N. Trendovi u skrbi za osobe starije životne dobi u Republici Hrvatskoj i u zemljama Europske Unije //Revija za socijalnu politiku, Vol.15 No.2 Lipanj 2008. Available from: http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=47998 [accessed 2012-12-23].

[3] Pečjak, V. Psihologija treće životne dobi, Zagreb: Naklada Prosvjeta, 2001, page 248.

[4] Mental Health Foundation. What is social inclusion and why is it so important? Like minds, line mine 2007; 30.
Fokkema T, Knipscheer K. Escape loneliness by going digital: A quantitative and qualitative evaluation of a Dutch experiment in using ECT to overcome loneliness among older adults. Aging & Mental Health 2007; 11(5): 496-504.

[5] Hernandez-Encuentra E, Pousada M, Gomez-Zuniga B. ICT and older people: beyond usability. Educational Gerontology 2009; 35: 226-245.

[6] Nahm E-S, Resnick B, Gaines J. Testing the reliability and validity of computer-mediated social support measures among older adults. CIN: Computers, Informatics, Nursing 2004; 22 (4): 211-219.

Project partners

The project *"Students for seniors: Basic ICT education for the elderly"* used a unique service learning approach to the ICT education for the elderly. It stretched over a semester and took place in the Tin Ujević public library and Trešnjevka retirement home.

Tin Ujević public library is part of a large library system *The Zagreb City Libraries*. Their mission is to ensure that all the citizens of Zagreb have free and equal access to all types of information, wide spectre of knowledge, national and world's heritage and diverse forms of cultural happenings, as well as to the basics of lifelong learning.

Trešnjevka retirement home in Zagreb provides care for the elderly population. They organize various activities and classes to raise the quality of life of their residents.

The library was chosen as a partner in this project because it is a public institution known for its accessibility, openness, and the initiative of the lifelong learning concept. The partnership with the library was established to attract all the home residents who wish to improve their ICT skills, but also to raise public awareness about a library as a place eager to satisfy the information, entertainment and educational needs of all social groups.

The retirement home was involved in the project with the intention to bring in and motivate as many members of the elderly population as possible, both the library users and those who do not necessarily recognize the advantage of the library services, so they could gain new ICT skills by participating in the project.

Project planning and realization

The project *"Students for seniors: Basic ICT education for the elderly"* was based around four workshops, each lasting for three hours having the maximum of eight participants per workshop. The project was, among other reasons, created as a support to the initiative of aging in an active, healthy way, which includes adapting oneself to the new circumstances, acquiring information and communications skills and lifelong education.

Also, it was important for students to introduce the elderly to the positive aspects of the Internet and present them with the ways in which they can benefit from it, bringing them closer to all the advantages ICT has to offer, even if it is only for a brief time.

Around 15 people, with different levels of knowledge partook in our information literacy workshops. The group, consisting mostly of female members,

Karahasanović A, Brandtzæg PB, Heim J, Lőders M, Vermeir L, Pierson J, Lievens B, Vanattenhoven J, Jans G. Co-creation and user-generated content – elderly people's user requirements. Computers in Human Behavior 2009; 25: 655-678.

was divided into beginner and the advanced users, who got a chance to reiterate what they had previously learned.
Students worked hard to motivate the elderly users to enrol in the workshop and to complete it successfully, in order to raise their interest to continue improving their skills, breaking some social stereotypes mentioned in the introductory part of this paper. None of the users showed up for the first workshop; it took a lot of convincing and additional motivation from the students to raise the attendance of their workshops.
Another important issue students had to consider was the physical state of the elderly. Although many members of the elderly population in the retirement home were happy to have a chance to participate in the workshops, they were physically limited having severe disabilities which prevented them from participating (such as hearing loss, poor eyesight, arthritis, limb or full body paralysis due to a stroke, etc). Unfortunately, these members could not have been taken into consideration as the participants in the workshops.

Project results

A total of four workshops were held (on November 28th, December 5th, 12th and 19th of 2012). Workshops had 15 attendants, out of which the most were women (more specifically, there were 14 female participants and only one male attendant).
The participants were encouraged to recognize the advantages information and communication technology has to offer, to learn how to implement the ICT in their daily routine, raising the quality of their life (using social networks or freely available online newspapers).
Workshop topics were divided into 4 groups of ICT skills:

1. Basic computer skills (starting up and turning off the computer, understanding and executing essential functions of the mouse, the keyboard, windows, simple text messages on the screen and basic functions in a browser, managing browser windows and browsing to specified pages using search engines, freely available online newspapers)
2. Email (starting up an email application, writing and receiving emails)
3. Creating and managing Facebook profile
4. Creating and managing Skype profile

All participants in the workshops were expected to acquire all skills from the 4 groups of ICT skills described above. The adoption of skills from the previous groups was a prerequisite for the following group.
Restrictions have been such that, depending on the mental and physical abilities of attendants, their attention capabilities and the time limit of the workshop, participants progressed in different rhythms. Some of them were not able to work through all the levels. Out of 15 participants, the first level (basic computer skills) was successfully completed by 13 participants. Two participants had physical deficiencies, shaking their hands and having poor eyesight, so they

could not coordinate the movements of their body (hands) with the visual input (display on the computer screen). The second level was completed by 9 participants. Attendants who were not able to complete this level were limited in time (they lost too much time mastering the computer basics).

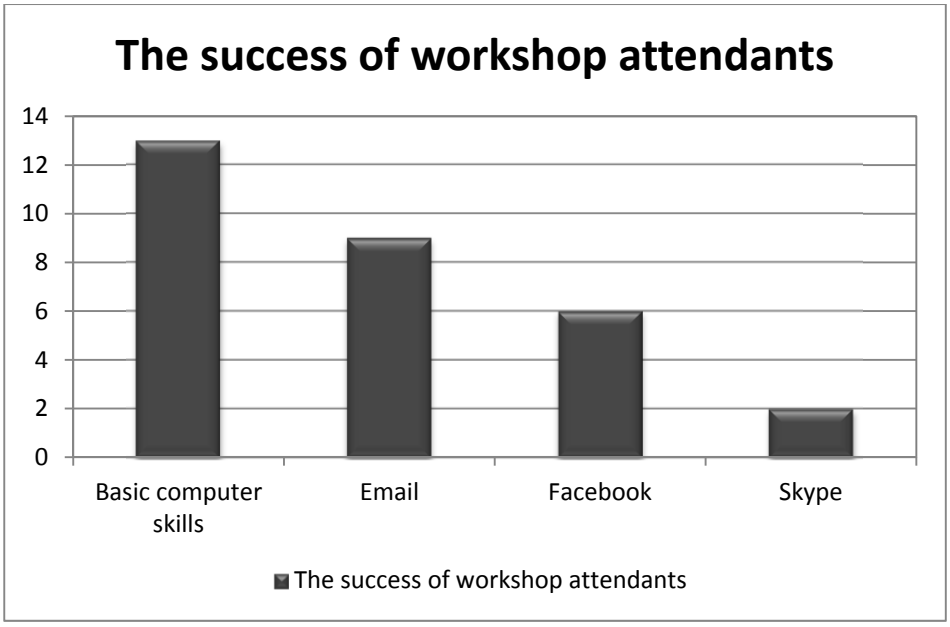


Figure 1. The performance of elderly using ICT technologies

Furthermore, 6 participants were able to complete the third level. Those who were not able to successfully adopt skills related to creating and managing the Facebook profile had trouble understanding all the options that Facebook offers; these skills were too complex compared to the time they had left after mastering the first two levels.

The last level (creating and managing the Skype account) was successfully completed by two participants. One of the reasons for such a low success rate was technical in nature (the majority of students did not have a camera installed on their home computer and therefore had no interest in this voice-over-IP service and instant messaging client). The other reason was a time limit. These two students who completed the final level and became familiar with Skype in a short time had a high motivation to acquire these skills.

Discussion

These workshops provided the basic grounding for the further use of the information technology in the retirement home, motivating the end users to continue using their ICT knowledge and build upon it.

The workshops were adapted to the needs of the elderly through an interactive teaching and learning process. Students were not lecturing, but rather guiding and pointing elderly in the right direction, being very patient and not rushing the learning process. It was crucial for elderly not to feel forced upon at any point of the workshop[7]. Beginners also needed help to overcome the fear of computers. The elderly were constantly encouraged and animated by our students to improve on the existing knowledge they have.

Considering that the main preference of the elderly in their daily routine is to read newspapers, and since many of them cannot afford buying newspapers on a daily basis, having access and skills to online versions of their preferred journals and newspapers made our attendants feel less excluded from the contemporary digital age and more informed about the local and global events.

It is especially important to emphasize the huge role that the library played in the entire process, as a social centre where the elderly were able to expand their horizons, improve their social relations and meet new people with similar interests and hobbies using the Internet, as well as to communicate through social networks with their family, friends in different countries, etc.

The mission of the student service learning project was achieved: the elderly participants have learned basic ICT skills, were acquainted with computer equipment, they learned how to browse the Internet and search for different information.

Furthermore, this service learning project helped both university students and their faculty mentor to better understand the process of aging, the needs and interests of the elderly in our local society, especially information and communication needs of the third-age population.

Finally, since Trešnjevka retirement home owns ten laptops, the residents of the home (our end users) can continue to actively use the knowledge they acquired during the workshops.

Service learning project and the long-term changes in the society

This part of the paper will address all benefits of the project that were obtained by each of the actors.

The improvement of the collaboration between two institutions, Trešnjevka retirement home and Tin Ujević public library is one of the many positive outcomes of this service learning project. After the workshops have finished, more

[7] Maravić, J. Cjeloživotno učenje. // Časopis Edupoint 3, 17(2003). Available from: http://www.carnet.hr/casopis/17/clanci/5 [accessed 2013-01-07].

elderly people came to the library on a daily basis, wanting to acquire new information and knowledge.
The other positive after-effect of the project is that students, their mentor, retirement home staff and library staff recognized the growing need for the continuous process of developing information literacy of the elderly, their introduction to the world of computers, the Internet and social networks.
Students have also concluded that the feeling of satisfaction for the exceptionally positive feedback from the elderly participants (who commented on how they liked the workshops, how useful they found them and how they helped them to learn many useful things), is priceless.
Creating a quality student- elderly relationship, developing and showing mutual respect, was another positive consequence of this project. During the workshops, the students have strived to develop a good relationship with the elderly, full of respect and patience.
The elderly showed that they felt it and reciprocated. While the elderly participants were happy and pleased they gained new skills, the students felt that their work was not purposeless because of the positive feedback. As a result, the students learned to teach, to be patient, to work in a group setting with the elderly and understand their learning pattern and the process of acquiring new knowledge.

Conclusion

The active and healthy aging includes adapting to the new circumstances, acquiring information skills and lifelong learning so that the elderly population does not feel excluded from the trends of the modern society. Therefore, the basic aim of the student service learning project *„Students for seniors: Basic ICT education for the elderly"* was to propagate the above mentioned actions. Elderly users who participated in the workshop lived in institutional care (homes for the elderly) and generally had fewer options of socializing. For these people, the computer presented a way of communicating with friends and family and become more active and involved in the society.
Apart from helping the elderly, the project enabled the students to gain the knowledge and the experience needed to work with a specific age group, but it also improved their communication skills, as well as the public speaking and the presentation skills.
The project was realized in the form of small, user-oriented workshops for the elderly. It was thoroughly planned over a course of semester taking into consideration the specific age, physical ability and "limitations" of the end users in acquiring knowledge and keeping their focus during the workshop. The project proved to be useful for both the personal and professional development of our end users. The skills they acquired are applicable in their everyday life and might enrich their daily activities and fill their time.

The greatest challenge for students was to adapt the teaching approach and presentation of information to a group of people in which the majority of the members worked with this type of technology for the first time, and to explain the basic information and communication terminology in a way that is most appropriate to them. The elderly have not only mastered information literacy, but have learned to use the computer for their everyday activities and to communicate with their acquaintances through social networks.

Finally, this service learning project proved that even the basic ICT education can enrich the lives of elderly and positively influence their social inclusion by expanding their personal social networks.

References

Fokkema T, Knipscheer K. Escape loneliness by going digital: A quantitative and qualitative evaluation of a Dutch experiment in using ECT to overcome loneliness among older adults. Aging & Mental Health 2007; 11(5): 496-504.

Hernandez-Encuentra E, Pousada M, Gomez-Zuniga B. ICT and older people: beyond usability. Educational Gerontology 2009; 35: 226-245.

Karahasanović A, Brandtzæg PB, Heim J, Lőders M, Vermeir L, Pierson J, Lievens B, Vanattenhoven J, Jans G. Co-creation and user-generated content – elderly people’s user requirements. Computers in Human Behavior 2009; 25: 655-678.

Laklija, M.; Rusac, S.; Žganec, N. Trendovi u skrbi za osobe starije životne dobi u Republici Hrvatskoj i u zemljama Europske Unije //Revija za socijalnu politiku, Vol.15 No.2 Lipanj 2008. Available from: http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=47998 [accessed 2012-12-23].

Lewis S, Ariyachandra T. Seniors and online Social network use. CONISAR 2010: 1-15.

Maravić, J. Cjeloživotno učenje. // Časopis Edupoint 3, 17(2003). Available from: http://www.carnet.hr/casopis/17/clanci/5 [accessed 2013-01-07].

Mental Health Foundation. What is social inclusion and why is it so important? Like minds, line mine 2007; 30.

Nahm E-S, Resnick B, Gaines J. Testing the reliability and validity of computer-mediated social support measures among older adults. CIN: Computers, Informatics, Nursing 2004; 22 (4): 211-219.

Pečjak, V. Psihologija treće životne dobi, Zagreb: Naklada Prosvjeta, 2001.

E-learning from the Perspective of Teachers at the University of Zagreb

Anamaria Kudrna
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
akudrna@ffzg.hr

Lana Ljevaković
Department of Information and Communication Sciences,
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, Zagreb, Croatia
lljevako@ffzg.hr

Summary

The purpose of this article is to investigate teachers' attitude about e-learning; it attempts to investigate the current teachers' opinions and awareness about e-learning, their activities pertaining to e-learning, as well as their motivation and initiative, or lack of them. The target group were the teaching staff of the faculties of the University of Zagreb who were sent a link to the online survey. By looking at the results of the University Computing Centre's annual survey, which was sent to University of Zagreb's faculty administration, it was assumed that teachers show motivation and interest in using e-learning, however, they lack support from their respective faculties, as well as technological prerequisites to implement e-learning in their teaching. Another assumption was that some faculties would have bigger and some smaller need for e-learning, depending on their field of study. To test these hypotheses, a survey was sent to each of the University's faculty with instructions to pass it along to members of their teaching staff. There were 315 participants in the survey, and data analysis confirmed on of the authors' hypotheses; it was concluded that the majority of participant teaching staff at the University, irrespective of their age, is motivated and interested in the idea of e-learning, but lack the incentive and support to engage with it.

Key words: e-learning, teacher motivation, external incentives, motivation, attitude, ambition, University of Zagreb, University Computing Centre, SRCE

Introduction

The existence and importance of e-learning is becoming harder to ignore as we witness the rapid improvement of ICT and the widespread presence of internet, and due to this, e-learning is becoming an increasing and inevitable trend in education. According to the University of Zagreb's Strategy on E-learning, it is not an alternative to standard education, but an integral part of it, an enhancement even [4]. The University Computing Centre (SRCE), conducts an annual survey on the subject of e-learning at the University of Zagreb [1]. This survey was an important orientational mark in this research and showed that the next logical step in exploring the phenomenon of e-learning is to investigate teachers' attitude towards it. From the survey it can be concluded that there are four key elements which are of utmost importance to e-learning: technological framework, attitude and support of the faculty administration, teachers' motivation, and active participation of the student body. Since the Centre's survey focused on the attitudes of faculty administrations, analyzing the technological framework was outside of our interest domain, and investigation of the opinions of the student body was already conducted at the Faculty of Humanities and Social Sciences, where most of the students expressed high interest in e-learning [3], the goal of this research was to explore attitudes and motivation of teachers, as well as some of their habits pertaining to implementation of e-resources and e-learning. The research didn't delve into the increase of teachers' responsibilities caused by e-learning; the focus was on their inclination to implement e-learning and their attitude towards their faculties' incentives for e-learning.

Hypothesis

The hypothesis of this research was drawn from the data collected by the Centre through their annual survey on e-learning at the University of Zagreb. One of the conclusions from that survey was that attitude of the faculty administration somewhat affects the attitude and motivation of teachers towards e-learning. This was also mentioned in notes that participants of the survey left: they gave a suggestion that it would be highly motivating if the University started to value participation in e-learning more. Therefore, the assumption of this hypothesis is that teachers lack incentive in implementation of e-learning into their courses. It was also assumed that teachers at technical and natural sciences faculties would be more prone to using e-learning than teachers in the social and humanistic field of study.

Method

The survey was created in Google Form [2] and was sent via e-mail to Vice Deans of Academic Affairs and representatives for e-learning from each faculty at the University of Zagreb. In some cases, it was also sent to each teacher personally. Structurally divided into four categories, the survey consists of twenty

questions. The first category helped the authors to create a demographic profile of the participant. The second group of questions investigates e-learning activities of teachers and their awareness of e-learning systems and incentives on the University level. The third group of questions focuses on their opinion about existing incentives for e-learning offered by their faculties. The fourth category pertains to the teachers' ambition, motivation and attitude towards using e-learning. The data was collected through the survey and organised in a spreadsheet, after which it was sorted and filtered. It was also interpreted by calculating and observing correlations between certain groups of data.

Results

The survey was filled by 315 participants, 56% of which were women. The oldest participant was 75, while the youngest was 23 years old. In Figure 1, the number of participants who use e-learning is analogous to the number of participants who don't use e-learning. This means that the tendency to use e-learning and participants' age are not interdependent.

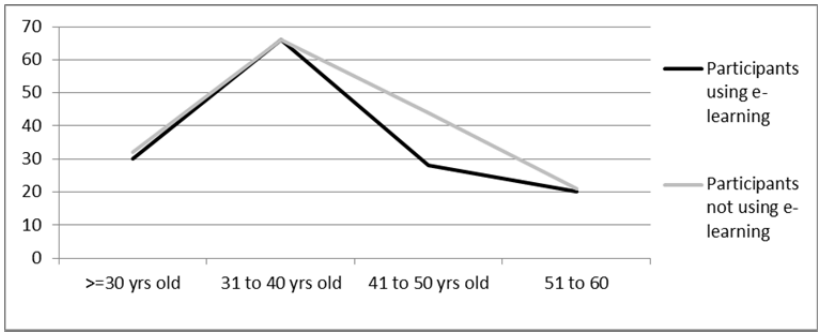


Figure 1. Correlation between age and tendency to use e-learning

The largest turnout was from the Faculty of Humanities and Social Sciences, whose staff comprise 21% of the participants of the survey. The survey was filled mostly by assistants, to be exact; 30%. Most of the participants; 99%, use the computer during their preparation for class, while 88% of participants also use the computer in teaching. 84% of participants answered that they know what systems for e-learning and learning management systems are, however, only slightly more than half of the participants; 53%, answered that they don't use any kind of e-learning system in teaching. The majority of participants; 85%, would consider it useful to track student online course activity on e-learning systems. Participants who do use an e-learning system use Moodle the most – 45.1%. Most of the participants; 82%, answered that they are aware that the University Computing Centre annually invites participants to apply for the best e-learning course award. Furthermore, 74% of participants know that the University of Zagreb gives out awards in the same category. 77% of participants

answered that they were informed about e-learning by their faculty, however, most of the participants; 88%, never received any incentive from their faculty to use e-learning. In fact, 78% of participants claim that their faculties don’t require of them to use an e-learning system. Figures 2 and 3 show that there are certain correlations within groups of participants who were informed by their faculties and participants who are required to use e-learning to groups of participants who answered that they use, or don’t use e-learning. The figures suggest that more teachers use e-learning in the cases where they were informed about e-learning, and they also suggest that more teachers use e-learning in the cases where their faculty requires its implementation.

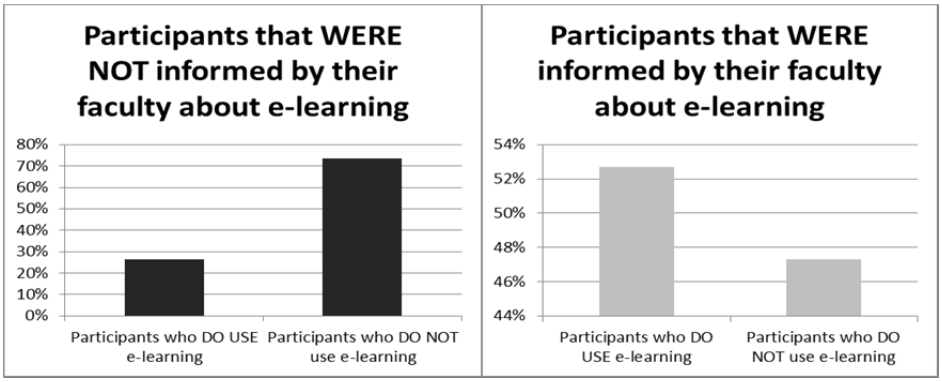


Figure 2. Correlation between participants’ e-learning usage and their faculty’s initiative to inform them about e-learning

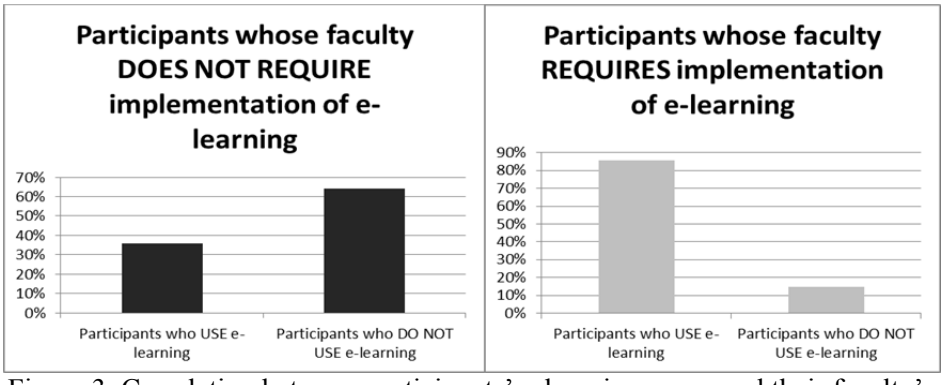


Figure 3. Correlation between participants’ e-learning usage and their faculty’s requirement regarding implementation of e-learning

Further; only 8% of the participants think that their faculty dedicates a lot of funds to the implementation of e-learning systems, while 32% think the amount of funds is small. 7% believe there are no funds allocated for implementation of e-learning at their faculty, and 7% don’t know whether there are any funds allo-

cated towards e-learning. Most of the participant; 61%, don't think that e-learning should be added as one of the criteria for career advancement. To the question whether they were satisfied with the incentive for implementation of e-learning that they get from their faculty, 22% answered with not satisfied at all, 10% said they were completely satisfied, while 18% answered they didn't know about such incentives. Participants were also asked to grade their motivation to be educated about e-learning. Their answers can be seen in Figure 4.

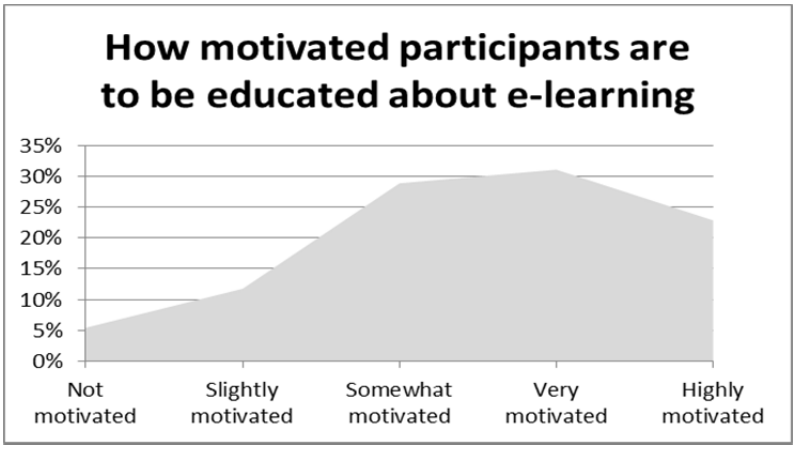


Figure 4. Participants' motivation to be educated about e-learning

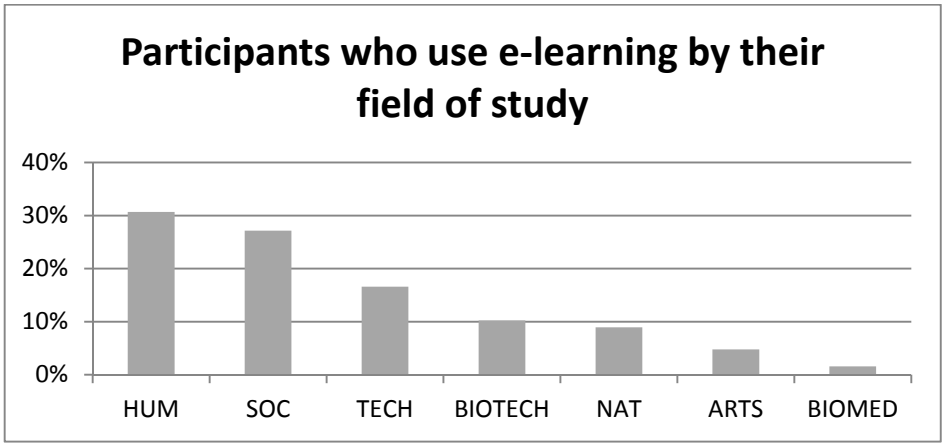


Figure 5. Correlation between participants' field of study and their e-learning usage

Furthermore, most of the participants; 61% of them, said they would be willing to set aside three hours a week during the course of several months to learn about e-education. Next, Figure 5 shows fields of studies of participants who

answered that they use e-learning. From the most to least represented, the field of studies are as follows: humanities (HUM), social sciences (SOC), technical sciences (TECH), biotechnical sciences (BIOTECH), natural sciences (NAT), arts (ARTS), and biomedicine (BIOMED). Data shown in the figure refutes the authors' hypothesis that teachers from technical and natural fields of study would be more prone to using e-learning than teachers in the social and humanistic fields of study, since it proves the opposite.
Furthermore, as shown in Figure 6, most of the participant teacher staff has a positive attitude about improving the quality of their courses by implementing e-learning.

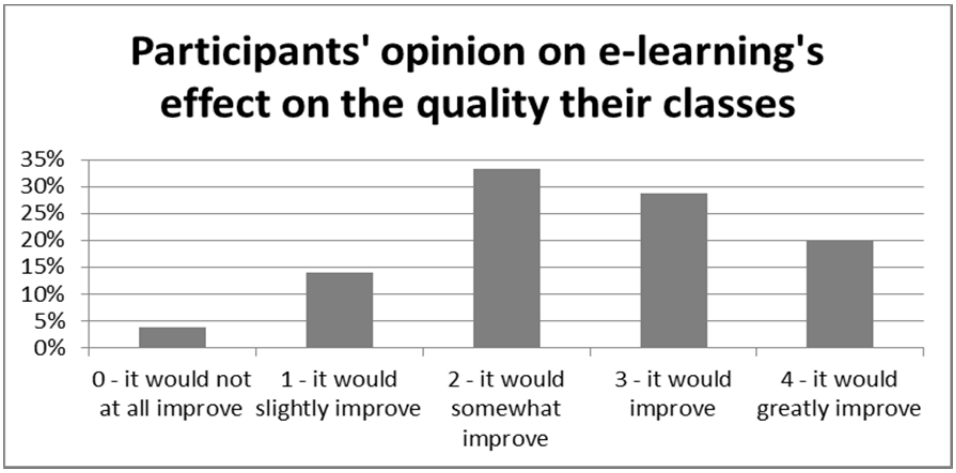


Figure 6. Participants' opinion on the effect of e-learning on the quality of classes

Conclusion

By analysing the data, it can be seen that there is no major difference in age between participants who answered that they use e-learning and participants who don't use e-learning. This suggests that age isn't a major factor in tendency to use e-learning. However, the hypothesis that teachers from technical and natural fields of study would be more prone to using e-learning was refuted. Most of the participants believe that e-learning leaves a positive effect on the quality of education. Greater number of participants also expressed that they would be motivated to further educate themselves on e-learning. The data also suggests that being informed by the faculty is a moderately influencing factor in e-learning usage. Other data shows that there is correlation between the faculty's policy on e-learning and teachers' e-learning activity in the sense that faculty policy seems to encourage teachers to use e-learning.

References

[1] Centar za e-učenje. Rezultati ankete o e-učenju na Sveučilištu u Zagrebu (prosinac 2012. godine). May 2013. http://www.unizg.hr/fileadmin/rektorat/dokumenti/e-ucenje/Sveuciliste_u_Zagrebu_Anketa_e_ucenje_2012_Rezultati_20130604.pdf (30.08.2013.)

[2] GoogleForm. https://support.google.com/drive/topic/1360904?hl=en&ref_topic=2811744 (May 2013)

[3] Pavešić, Leon. Istraživanje mišljenja studenata Filozofskog fakulteta u Zagrebu o Omega sustavu za učenje na daljinu. Zagreb, 2012

[4] University of Zagreb. E-learning strategy 2007-2010. May 2007. http://www.unizg.hr/fileadmin/rektorat/dokumenti/eucenje_strategija/University_of_Zagreb-E-learning_strategy.pdf (30.08.2013.)

INFuture Conference Standing Board of Reviewers

Vjekoslav Afrić, Faculty of Humanities and Social Sciences, University of Zagreb
Winton Afrić, Faculty of Teacher Education, University of Zagreb
Željko Agić, Faculty of Humanities and Social Sciences, University of Zagreb
Zvjezdana Antoš, Etnographic Museum Zagreb
John Akeroyd, South Bank University, London
Yvette M. Alex-Assensoh, Indiana University Bloomington
Silvija Babić, Croatian State Archives, Zagreb
Miroslav Bača, Faculty of Organization and Informatics, University of Zagreb Varaždin
Mihaela Banek Zorica, Faculty of Humanities and Social Sciences, University of Zagreb
Ana Barbarić, Faculty of Humanities and Social Sciences, University of Zagreb
Anabela Barreiro, Institute for Systems and Computer Engineering: Research and Development – INESC-ID, Lisbon
Bob Bater, InfoPlex Associates Ltd., Bristol
David Bawden, City University, London
Domagoj Bebić, Faculty of Political Science, University of Zagreb
Zoran Bekić, University Computing Center, Zagreb
Inoslav Bešker, University of Bologna
Clive Billenness, Best Practice User Group, Leeds, United Kingdom
Helen Boelens, The Netherlands
Maurice Brandt, Research Data Centre, Federal Statistical Office Germany
Goran Bubaš, Faculty of Organization and Informatics, University of Zagreb Varaždin
Viktorija Car, Faculty of Political Science, University of Zagreb
Andrew Cox, University of Sheffield
Alma Leora Culén, Department of Informatics, University of Oslo
Oliver Čulo, Universität Mainz
Blaženka Divjak, Faculty of Organization and Informatics, University of Zagreb
Senada Dizdar, Faculty of Philosophy, Sarajevo
Milena Dobreva, University of Portsmouth
Zdravko Dovedan Han, Faculty of Humanities and Social Sciences, University of Zagreb
Zvjezdana Dukić, Hong Kong
Annette Đurović, Faculty of Philology, University of Belgrade
Karmen Erjavec, Faculty of Social Sciences, University of Ljubljana
Aleš Gačnik, Znanstveno-raziskovalno središče Bistra, Ptuj
Franz Guenthner, Centrum für Informations- und Sprachverarbaitung, Ludwig-Maximilians-Universität München

Ivana Hebrang Grgić, Faculty of Humanities and Social Sciences, University of Zagreb
Anette Hemera, CSoft International, Budapest
Eulàlia Hernández-Encuentra, Open University of Catalonia
Aleksandra Horvat, Faculty of Humanities and Social Sciences, University of Zagreb
Maja Jokić, Institute for Social Research, Zagreb
Vedran Juričić, Faculty of Humanities and Social Sciences, University of Zagreb
Jelena Jurišić, Croatian Studies, University of Zagreb
Igor Kanižaj, Faculty of Political Science, University of Zagreb
Kristina Kocijan, Faculty of Humanities and Social Sciences, University of Zagreb
Božidar Kovačić, Department of Informatics, University of Rijeka
Lejla Kodrić Zaimović, Faculty of Philosophy, Sarajevo
Steven Krauwer, Utrecht institute of Linguistics UiL-OTS, The Netherlands
Vlasta Kučiš, University of Maribor
Serap Kurbanoglu, Department of Information Management, Hacettepe University, Ankara
Stefan Langer, Centrum für Informations- und Sprachverarbaitung, Ludwig-Maximilians-Universität München
Jadranka Lasić-Lazić, Faculty of Humanities and Social Sciences, University of Zagreb
Tomislava Lauc, Faculty of Humanities and Social Sciences, University of Zagreb
Nikolaj Lazić, Faculty of Humanities and Social Sciences, University of Zagreb
Marko Lukičić, Corporate Strategy and Business Innovation Sector, Hrvatski Telekom, Zagreb
Maurizio Lunghi, Fondazione Rinascimento Digitale, Florence
Nikola Ljubešić, Faculty of Humanities and Social Sciences, University of Zagreb
Stjepan Malović, University of Dubrovnik
Luisa Marquardt, Roma Tre University
Vladimir Mateljan, Faculty of Humanities and Social Sciences, University of Zagreb
Nives Mikelić Preradović, Faculty of Humanities and Social Sciences, University of Zagreb
Anja Nikolić Hoyt, Croatian Academy of Sciences and Arts
Zoran Nježić, Faculty of Teacher Education, University of Zagreb
Đorđe Obradović, University of Dubrovnik
Aleksander Panjek, University of Primorska, Koper
Krešimir Pavlina, Faculty of Humanities and Social Sciences, University of Zagreb

Jelka Petrak, School of Medicine, University of University of Zagreb
Ana Pongrac Pavlina, Faculty of Humanities and Social Sciences, University of Zagreb
Nenad Prelog, University of Zadar
Željka Radovinović, Zagreb Academy of music
Arian Rajh, Agency for Medical Products and Medical Devices Zagreb
Andreas Rauber, Vienna University of Technology (TU-Wien)
Sanja Seljan, Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Simeon, Faculty of Humanities and Social Sciences, University of Zagreb
Karolj Skala, Ruđer Bošković Institute, University of Zagreb
Aida Slavić, UDC Consortium
Sonja Špiranec, Faculty of Humanities and Social Sciences, University of Zagreb
Hrvoje Stančić, Faculty of Humanities and Social Sciences, University of Zagreb
Marko Tadić, Faculty of Humanities and Social Sciences, University of Zagreb
Tomislav Tepeš, Emirates Airline, Dubai
Martin Tröndle, Zeppelin University, Friedrichshafen
Miroslav Tuđman, Faculty of Humanities and Social Sciences, University of Zagreb
Lejla Turčilo, Faculty of Political Science, University of Sarajevo
Ludmila Ticha, Czech Technical University in Prague
Kostandina Veljanovska, Faculty of Administration and Managment of Information Systems, St. Kliment Ohridski University, Bitola
Radovan Vrana, Faculty of Humanities and Social Sciences, University of Zagreb
Neven Vrček, Faculty of Organization and Informatics, University of Zagreb Varaždin
Aleksandra Vraneš, Faculty of Philology, Belgrade University
Žarka Vujić, Faculty of Humanities and Social Sciences, University of Zagreb
Mirna Willer, Department of Library and Information Science, University of Zadar
Goran Zlodi, Faculty of Humanities and Social Sciences, University of Zagreb
Vatroslav Zovko, Faculty of Teacher Education, University of Zagreb
Daniela Živković, Faculty of Humanities and Social Sciences, University of Zagreb

The INFuture2013 conference was supported by

Goethe Institut
http://www.goethe.de/ins/hr/zag/hrindex.htm

Project XLIKE
http://www.xlike.org/

Zagreb Tourist Board
http://www.zagreb-touristinfo.hr/

Media coverage of the INFuture2013 conference

Infotrend
http://www.infotrend.hr/