

An International Journal on Grey Literature



Autumn 2014 – TGJ Volume 10, Number 3

‘WEIGHING UP PUBLIC ACCESS TO GREY LITERATURE’

GreyNet

The Grey Journal

An International Journal on Grey Literature

COLOPHON

Journal Editor:

Dr. Dominic J. Farace
Grey Literature Network Service
GreyNet, The Netherlands
journal@grey.net

Associate Editors:

Julia Gelfand
University of California, Irvine
UCI, United States

Dr. Debbie L. Rabina
School of Information and Library Science
Pratt Institute, United States

Dr. Joachim Schöpfel
Université Charles de Gaulle Lille 3
France

Gretta E. Siegel, Professor Emeritus
Portland State University
PSU, United States

Kate Wittenberg
Project Director, Client and Partnership
Development at Ithaka, United States

Technical Editor:

Jerry Frantzen, TextRelease

About TGJ

The Grey Journal is a flagship journal for the international grey literature community. It crosses continents, disciplines, and sectors both public and private.

The Grey Journal not only deals with the topic of grey literature but is itself a document type classified as grey literature. It is akin to other grey serial publications, such as conference proceedings, reports, working papers, etc.

The Grey Journal is geared to Colleges and Schools of Library and Information Studies, as well as, information professionals, who produce, publish, process, manage, disseminate, and use grey literature e.g. researchers, editors, librarians, documentalists, archivists, journalists, intermediaries, etc.



CIP

The Grey Journal (TGJ): An international journal on grey literature / D.J. Farace (Journal Editor); J. Frantzen (Technical Editor) ; GreyNet, Grey Literature Network Service. - Amsterdam: TextRelease, Volume 10, Number 3, Autumn 2014. - *CVTISR, EBSCO, FEDLINK-Library of Congress, Inist-CNRS, ISTI-CNR, KISTI, NIS-IAEA, NTK, and NYAM are Corporate Authors and Associate Members of GreyNet International. This serial publication appears three times a year - in spring, summer, and autumn. Each issue is thematic and deals with one or more related topics in the field of grey literature. The Grey Journal appears both in print and electronic formats.*
ISSN 1574-1796 (Print)
ISSN 1574-180X (E-Print/PDF)
ISSN 1574-9320 (CD-Rom)

Subscription Rates:

€240 institutional, including postage & handling

Contact Address:

Reprint Service, Back Issues, Document Delivery,
Advertising, Inserts, Subscriptions:

TextRelease
Javastraat 194-HS
1095 CP Amsterdam
Netherlands
T/F +31 (0) 20 331.2420
info@textrelease.com
<http://www.textrelease.com/glpublishings.html>

About GreyNet

The Grey Literature Network Services was established in order to facilitate dialog, research, and communication between persons and organizations in the field of grey literature. GreyNet further seeks to identify and distribute information on and about grey literature in networked environments. Its main activities include the International Conference Series on Grey Literature, the creation and maintenance of web-based resources, a moderated Listserv, and The Grey Journal. GreyNet is also engaged in the development of distance learning courses for graduate and post-graduate students, as well as workshops and seminars for practitioners.

Full-Text License Agreement

In 2004, TextRelease entered into an electronic licensing relationship with EBSCO Publishing, the world's most prolific aggregator of full text journals, magazines and other sources. The full text of articles in The Grey Journal (TGJ) can be found in *Library, Information Science & Technology Abstracts (LISTA)* full-text database.

© 2014 TextRelease

Copyright, all rights reserved. No part of this publication may be reproduced, stored in or introduced into a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without prior permission of the publisher.

Contents

'WEIGHING UP PUBLIC ACCESS TO GREY LITERATURE'

Using Google Search Appliance (GSA) to Search Digital Library Collections : A Case Study of the INIS Collection Search	123
<i>Dobrica Savić (Austria)</i>	
Grey Literature in European Commission Projects	133
<i>Sara Goggi, Gabriella Pardelli, Silvia Giannini and Stefania Biagioni (Italy)</i>	
Effect of Open API, NDSL Open Service (NOS) on Sharing Technical Reports in Korea	145
<i>Seon-Hee Lee and Mi Hwan Hyun (Korea)</i>	
Effects of legal changes and electronic publishing on the access to GL in Iceland	151
<i>Stefanía Júlíusdóttir (Iceland)</i>	
Scholarly publishing behaviour in Slovakia ... Are we ready for repositories?	163
<i>Jaroslav Šušol, Miriam Ondrišová, Jana Ilavská, and Marcela Kopecká (Slovakia)</i>	
Back to Grey: Disclosure and Concealment of Electronic Theses and Dissertations	173
<i>Joachim Schöpfel and Hélène Prost (France)</i>	

Colophon	118
----------------	-----

Editor's Note	121
---------------------	-----

On the News Front

Investigation into the use of Grey Literature in Evidence-Based Medicine: The role of the NHS Library and Information Services Professional	182
<i>Joanna Hooper, United Kingdom</i>	

GL16 Call for Posters 'Grey Literature Lobby: Engines and Requesters for Change'	184
---	-----

GL16 Registration Sixteenth International Conference on Grey Literature, December 8-9, 2014, Library of Congress, Washington DC, USA	185
--	-----

Advertisements

EBSCO LISTA Full-Text	120
INIS, The International Nuclear Information System	122
Ila – Information International Associates, Inc.	172

About the Authors	186
-------------------------	-----

Notes for Contributors	187
------------------------------	-----

Library, Information Science & Technology AbstractsTM with Full Text

Available via EBSCOhost[®]

The definitive professional information
resource designed for librarians and
information specialists...

*Library, Information Science & Technology AbstractsTM
with Full Text* is an indispensable tool for librarians
looking to stay current in this rapidly evolving field.

Comprehensive content includes:

- Full text for more than 270 journals and nearly 20 monographs
- Indexing for more than 550 core journals, 50 priority journals and nearly 125 selective journals
- Includes books, research reports, proceedings and author profiles
- Access to 6,800 terms from reference thesauri
- Coverage extends back as far as the mid-1960s

Subject coverage includes:

- Bibliometrics
- Cataloging
- Classification
- Information Management
- Librarianship
- Online Information Retrieval
- And much more...

Contact EBSCO Publishing to learn more about
Library, Information Science & Technology AbstractsTM with Full Text,
or to request a free trial.

Phone: 800.653.2726

Email: request@ebscohost.com

www.ebscohost.com



Editor's Note

This issue of The Grey Journal marks the completion of its 10th volume, 2005-2014. The theme fittingly addresses the occasion "Weighing up Public Access to Grey Literature" or rendered in current political jargon as '*the optics of grey*'.

A decade ago, this flagship journal for the international grey literature community set course destined to cross continents, disciplines, and sectors of information both public and private. The Grey Journal with GreyNet International as its corporate author has over the years helped steer the frontend of research and education in the field of grey literature.

In recent months, a new change in course set out in the 'Pisa Declaration on Policy Development for Grey Literature Resources' has been drafted. This document is now open for online endorsement and will be ratified during a special session of the Sixteenth International Conference on Grey Literature held this December at the Library of Congress, Washington D.C.

In advance of its ratification, GreyNet seeks to comply with the Pisa Declaration and has synthesised the 15 points to five commitments in serving the international grey literature community:

First, the organizational commitment to open access, to further cooperation and coordination between and among grey literature communities, sharing open data standards.

Second, the commitment to research and education, where recognition and reward is associated with quality grey literature, and where attention is given to good practices in the field.

Third, the commitment to address and safeguard legal issues inherent to grey literature by exploring the various types of licensing agreements now available and by fostering constructive relations with commercial publishers.

Fourth, a commitment to sustainability linked to a financial prerequisite. Identifying funding and grants for special collections and repositories, commitment to long term preservation, and investments in new technologies.

Fifth, a firm technical commitment, a guarantee to continued online services and further crosslinking of textual and non-textual content. A commitment - from tackling broken links to facilitating interoperability - regardless of the system or portal in which grey literature and its accompanying data are housed.

And, to this end, The Grey Journal continues on course.

Dominic Farace
journal@grey.net.org

Grey Literature is a field in library and Information science that deals with the production, distribution, and access to multiple document types produced on all levels of government, academics, business, and industry in electronic and print formats not controlled by commercial publishing i.e. where publishing is not the primary activity of the producing body.

INIS

www.iaea.org/inis

The International Nuclear Information System

***The world's leading
source of nuclear
information since 1970***



IAEA

International Atomic Energy Agency

Using Google Search Appliance (GSA) to Search Digital Library Collections : A Case Study of the INIS Collection Search^{*}

Dobrica Savić (Austria)

Abstract

Google Search has established a new standard for information retrieval which did not exist with previous generations of library search facilities. The INIS hosts one of the world's largest collections of published information on the peaceful uses of nuclear science and technology. It offers on-line access to a unique collection of 3.7 million bibliographic records and 483,000 full texts of non-conventional (grey) literature. This large digital library collection suffered from most of the well-known shortcomings of the classic library catalogue. Searching was complex and complicated, it required training in Boolean logic, full-text searching was not an option, and response time was slow. An opportune moment to improve the system came with the retirement of the previous catalogue software and the adoption of GSA as an organization-wide search engine standard. INIS was quick to realize the potential of using such a well-known application to replace its on-line catalogue. This paper presents the advantages and disadvantages encountered during three years of GSA use. Based on specific INIS-based practice and experience, this paper also offers some guidelines on ways to improve classic collections of millions of bibliographic and full-text documents, while reaping multiple benefits, such as increased use, accessibility, usability, expandability and improving user search and retrieval experiences.

1 Introduction

The world of library, information and knowledge management is facing many challenges. Besides diminished funding and increased user expectations, the use of classic search tools, such as library catalogues, is becoming an additional stumbling block and imminent challenge. Library users require fast and easy access to full-text information resources, regardless of whether the original format is paper or electronic. Practice shows that users demand reliable and immediate access to the results of their queries, which should be performed through a simplified, non-complex user interface. However, most libraries still offer complex Boolean searches made through each individual library-specific interface, the results of which are displayed in a highly technical, standardized, professional library format, usually bearing no relevance to end users.

Google Search, with its huge coverage, high retrieval speed, and appealing simplicity, has established a new standard of information retrieval, which was not possible with the previous generation of library search solutions. Put in a position of David versus Goliath, many small, and even larger libraries, are losing the battle and letting many of its users use Google rather than library catalogues. Statistics back this up. In 2013, over two trillion searches were made using Google, or around six billion searches a day¹. At the same time, the Library of Congress, USA, one of the world's largest libraries, had a total of 2 billion searches since the establishment of its ILS in August 1999².

A question for the library community and practicing librarians is whether Google can be used to improve user satisfaction when visiting classic libraries, and if it can increase the number of visits and visitors. If so, then how can it be done? This was the question faced by the IAEA's International Nuclear Information System (INIS) when it decided to replace its aging search engine. In other words, the question and the challenge for INIS was how to revamp the on-line database catalogue with millions of bibliographic records and thousands of full-text nuclear documents, while increasing its use, accessibility, usability, and expandability. INIS, one of the world's largest collections of published information on the peaceful uses of nuclear science and technology, offers on-line access to a unique collection of 3.7 million bibliographic records and 483,000 full texts of non-conventional (grey) literature. However, searching the INIS Collection was complex and complicated process, and required training in using Boolean logic, while full-text searching was not an option, and response time was slow. An opportune moment came to upgrade the system with the retirement of the previous catalogue software and the adoption of Google Search Appliance (GSA) as an organization-wide search engine standard. INIS was quick to realize the potential of using such a well-known application as a replacement for its on-line catalogue. This paper presents the advantages and disadvantages, as well as some challenges encountered during the last three years of GSA use. Based on specific INIS practices, this paper shares experiences on ways to improve classic library catalogue, while reaping multiple benefits, namely increased use, better accessibility, easier usability, expandability and improved user search and retrieval experiences.

^{*} First published in J LIS.it. Vol. 5, n. 2 (Luglio/July 2014): Art: #10071. DOI: 10.4403/jlis.it-10071.Web
<http://leo.cineca.it/index.php/jlis/article/view/10071/9423>

2 IAEA Nuclear Information Section (NIS)

The IAEA is regarded as the world's centre of cooperation in the field of safe, secure and peaceful uses of nuclear technologies. It was set up in 1957 as the world's "Atoms for Peace"³ organization within the United Nations system. As of February 2014, the IAEA has 162 Member States.

The IAEA Secretariat is headquartered at the Vienna International Centre in Vienna, Austria. It also operates liaison and regional offices in Geneva, Switzerland; New York, USA; Toronto, Canada; and Tokyo, Japan. The IAEA runs or supports research centres and scientific laboratories in Vienna and Seibersdorf, Austria; Monaco; and Trieste, Italy. The IAEA Secretariat is a team of 2300 multi-disciplinary professional and support staff from more than 100 countries.



The IAEA's mission is guided by the interests and needs of Member States, strategic plans and the vision embodied in the IAEA Statute. Three main pillars — or areas of work — underpin the IAEA's mission: Safety and Security; Science and Technology; and Safeguards and Verification. The work of the IAEA is carried out through six departments⁴ — Nuclear Energy, Nuclear Safety and Security, Nuclear Science and Applications, Safeguards, Technical Cooperation, and the Department of Management. Although supporting the entire Agency, the NIS is organizationally part of the Department of Nuclear Energy. The Department's main tasks are to foster the efficient and safe use of nuclear power by supporting interested Member States in improving the performance of nuclear power plants, the nuclear fuel cycle, and the management of nuclear wastes; catalyzing innovation in nuclear power and fuel cycle technologies; development of indigenous capabilities for national energy planning; the deployment of new nuclear power plants; and the advancement of science and industry through improved operation of research reactors. One of its tasks is also the preservation and dissemination of nuclear information and knowledge, which in turn is the responsibility of Nuclear Information Section (NIS).

The Nuclear Information Section consists of the IAEA Library, the INIS Unit, and the Systems Development and Support Group. The INIS Unit fosters the collection and exchange of scientific and technical information on the peaceful use of nuclear science and technology; maintains a unique nuclear related thesaurus, increases awareness in Member States of the importance of maintaining efficient and effective systems for managing such information; provides information services and support to Member States and to the IAEA; and assists with capacity building and training.

3 International Nuclear Information System (INIS)

The Statute of the IAEA, Article III, states that the Agency is authorized to foster the exchange of STI on peaceful uses of atomic energy. Article VIII, which is devoted to the exchange of information, states that the Agency's goal is to foster the exchange of STI on the peaceful uses of atomic energy, to encourage the exchange among its members of information relating to the nature and peaceful uses of atomic energy, and that it shall serve as an intermediary among its members for this purpose. Based on these Statute provisions, INIS was created in 1969 to provide computerized access to an extensive collection of references to the world's nuclear literature. It was designed as an international cooperative venture, requiring the active participation of its members. While it started with only 25 members, today there are 153 (129 countries and 24 international organizations⁵). INIS also represents the world's first truly international computerized information system. Under the INIS concept, each participating member undertakes to look through literature published within its boundaries and select those documents that fall within the agreed subject scope. The countries prepare a detailed description of each item selected and send it, in some cases together with a copy of the document, to the INIS Secretariat in Vienna. Here, the incoming information is checked and combined with input from other countries into a single database collection.



INIS is a channel for information exchange that employs the very latest technology, thus, proving over the decades to be instrumental in bringing cutting edge technology to countries or geographical areas which lack such facilities or infrastructures. It is also the tool used by scientists, engineers, technicians, and managers in the nuclear industry to keep abreast of developments in the subject areas covered by the INIS Collection. From the perspective of "knowledge management and preservation", INIS is the repository for references to publications that contain cumulative scientific knowledge in the areas of peaceful applications of nuclear science and technology as recorded in scientific journals, as well as the repository for the full texts of grey literature, not easily available through regular commercial channels.

4 INIS Collection

INIS represents an extraordinary example of world cooperation, where 152 members allow access to their valuable nuclear information resources in order to preserve world peace and further increase the use of nuclear energy for peaceful purposes. There are more than 3.7 million bibliographic references to publications, documents, technical reports, non-copyrighted documentation, and other grey literature made available, but 483,000 full texts are also directly accessible. Overall, there are 800 GB of data in the INIS Collection. Besides being a source of information when searching, the availability of full-text gives INIS a special role — being the main custodian of this world information heritage and preserving this codified, specialized, scientific and technical knowledge. On average, INIS adds 120,000 bibliographic records and 13,000 full-text PDF documents to its collection annually. The complete collection is freely accessible from the INIS Search website⁶. The INIS Collection covers around 50 well defined subject categories which are regularly maintained by INIS and the ETDE. They also provide the scope descriptions used by national and regional centres to categorize nuclear literature for INIS input, and to categorize energy technology literature for ETDE input. The ETDE/INIS Joint Reference Series publications are also available on the INIS website⁷. The INIS Collection covers all aspects of the peaceful uses of nuclear science and technology such as nuclear reactors, reactor safety, nuclear fusion, applications of radiation and radioisotopes in medicine, agriculture, industry and pest control, as well as related fields of nuclear chemistry, nuclear physics and materials science. Special emphasis is placed on the environmental, economic and health effects of nuclear energy. Legal and social aspects associated with nuclear energy are also covered. Figure 1 lists a complete set of INIS Subject Categories. Figure 2 gives a break-down of the amount of documents by major category, while Figure 3 gives a similar break-down according to various document types in the Collection.

S01 – Coal, lignite, and peat	S42 – Engineering
S02 – Petroleum	S43 – Particle accelerators
S03 – Natural gas	S46 – Instrumentation related to nuclear science and technology
S04 – Oil shales and tar sands	S47 – Other instrumentation
S07 – Isotopes and radiation sources	S54 – Environmental sciences
S08 – Hydrogen	S58 – Geosciences
S09 – Biomass fuels	S60 – Applied life sciences
S10 – Synthetic fuels	S61 – Radiation protection and dosimetry
S11 – Nuclear fuel cycle and fuel materials	S62 – Radiology and nuclear medicine
S12 – Management of radioactive wastes, and non-radioactive wastes	S63 – Radiation, thermal, and other environmental pollutant effects on living organisms and biological materials
S13 – Hydro energy	S70 – Plasma physics and fusion technology
S14 – Solar energy	S71 – Classical and quantum mechanics, general physics
S15 – Geothermal energy	S72 – Physics of elementary particles and fields
S16 – Tidal and wave power	S73 – Nuclear physics and radiation physics
S17 – Wind energy	S74 – Atomic and molecular physics
S20 – Fossil fuel power plants	S75 – Condensed matter physics, superconductivity and superfluidity
S21 – Specific nuclear reactors and associated plants	S77 – Nanoscience and nanotechnology
S22 – General studies of nuclear reactors	S79 – Astrophysics, cosmology and astronomy
S24 – Power transmission and distribution	S96 – Knowledge management and preservation
S25 – Energy storage	S97 – Mathematical methods and computing
S29 – Energy planning, policy and economy	S98 – Nuclear disarmament, safeguards and physical protection
S30 – Direct energy conversion	S99 – General and miscellaneous
S32 – Energy conservation, consumption, and utilization	
S33 – Advanced propulsion systems	
S36 – Materials science	
S37 – Inorganic, organic, physical and analytical chemistry	
S38 – Radiation chemistry, radio chemistry and nuclear chemistry	

Figure 1: INIS Subject Categories

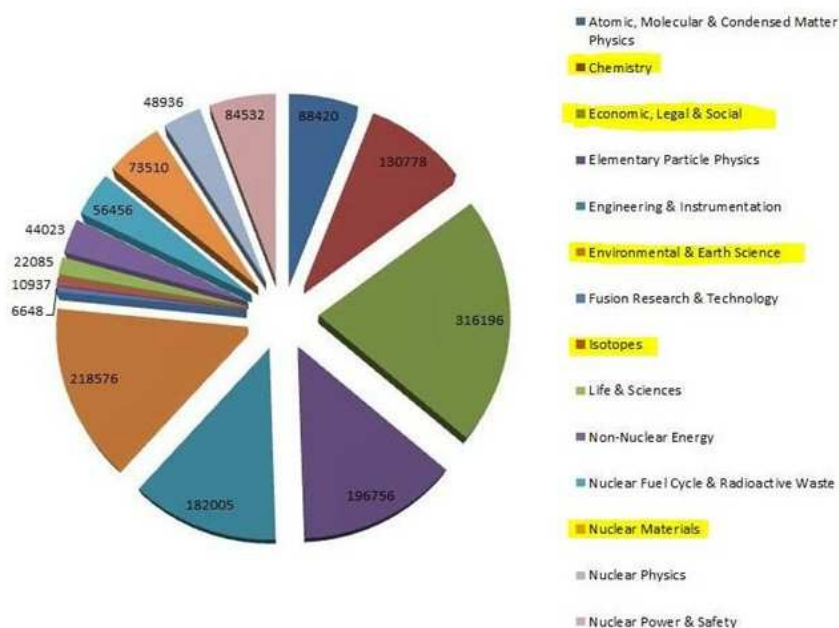


Figure 2: INIS Collection by Subject

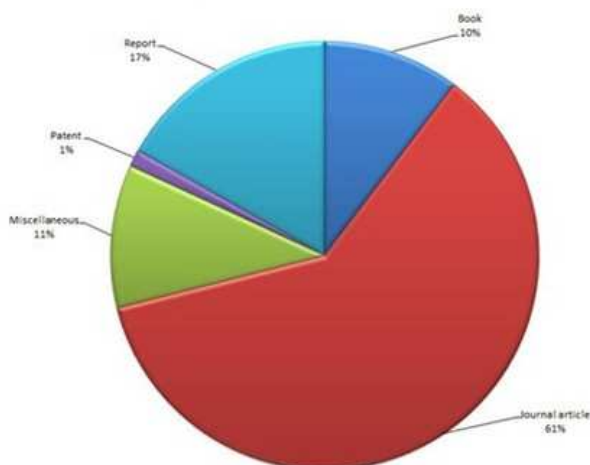


Figure 3: Bibliographic Records by Type (1 January 2014: 3,623,201 records)

5 INIS Search Engine

Since its inception, the INIS Collection has operated in a controlled environment, where users need to register through their national INIS centre, as well as the INIS Secretariat headquarters in Vienna, before being given access to the Collection. This changed in April 2009, when INIS became a free, open, and unrestricted information resource for internet users around the world. The opening of the Collection simplified access to reliable nuclear information on the peaceful uses of nuclear science and technology, including non-conventional literature, and made nuclear knowledge readily available worldwide for research, development and other uses. The opening of the Collection resulted in a significant increase of users. The old search engine, a well-known BASIS ERDMS⁸ was in operation from almost the beginning of INIS until 2011. Eventually, BASIS merged with an RDBMS system called DM, and became known as BASISplus. Currently, BASISplus is owned and supported by OpenText.

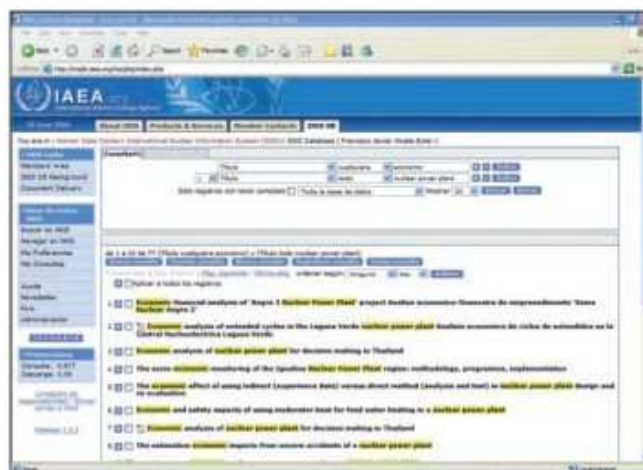


Figure 4: Old INIS BASIS Search Interface

The problems INIS experienced with its version of BASISplus included slow response time, a cluttered interface made for librarians, only an advanced search option, no full-text indexing, issues with support for a multilingual search interface, and lack of support for the thesaurus and different authorities. The new search engine was installed in 2010, and became fully operation in April 2011. It was based on GSA, which, at that time, became the IAEA-wide search engine standard. Although INIS reviewed a number of different search engine offerings, the decision was made to use GSA as a de facto on-line catalogue and to search the digital collection of INIS records.



Figure 5: ICS Standard Interfaces

This decision was based on the following characteristics offered by GSA: great speed and scalability; uncluttered and easy to use starting interface; the possibility to use advanced options and to broaden or tighten searches; powerful full-text indexing; retrieval of more relevant results; and the existence of faceted/filtered search features. The images shown above are actual snapshots of two different versions of the INIS GSA-based standard (simple) search interfaces, both deployed on the INIS website.

6 Main Features of the INIS Collection Search (ICS)

The INIS Collection Search I based on Google Search Appliance© technology⁹ — a renowned, simple, fast, and flexible search solution. Since its inception in 2010, the ICS has gone through a number of changes. Its current version, 4.3, was implemented in January 2014. The system is hosted internally on virtual servers within the IAEA and is constantly monitored for 24/7 availability. The INIS Collection Search, which in fact, is the INIS OPAC, includes the following main features:

6.1 Accessibility

ICS is an open web-based application freely accessible to Internet users interested in nuclear related publications. It's URL¹⁰, as well as the INIS general URL¹¹, are open and well known web addresses which are easy to find and locate using any search engine. In fact, when searching for "nuclear

information” on Google.com, the INIS home page is among the first results. Besides web access, INIS and ICS are available through mobile applications for iPad¹², iPhone¹³, and Android¹⁴. The NE News app, also available on the above devices, allows access to all of the IAEA Department of Nuclear Energy’s newsletters, brochures and social media channels through a single portal. This includes the authoritative Nuclear Energy Series of technical publications that cover a wealth of topics, ranging from introducing nuclear power to decommissioning. The same app provides access to the ICS.

A special widget — a compact version of the INIS Collection Search form that can be integrated into third-party Web sites — is also available. When users of third-party sites query the INIS Collection using the widget, they are redirected to the INIS website, where the search results are displayed as if the query had been launched directly from the INIS Collection Search. The ICS widget can be deployed with or without a search term filter. The Japan Atomic Energy Agency uses the former ICS search widget¹⁵. Examples of the latter widget are the ones installed on the IAEA Research Reactors website¹⁶, where the search is limited to research reactor documents, and the one on the IAEA Fast Reactor Technology website¹⁷.

6.2 Ease of Use

The Initial ICS screen (Figure 6) is intuitive and self-explanatory, enabling even those with no knowledge of specific INIS Collection characteristics to make queries and get desired results. It mirrors the Google.com user simplicity concept with one simple search box and very few additional, but optional commands. Based on this principle of straight-forward simplicity and usability, and due to the popularity of Google.com, previous knowledge is not needed to start using the system. The more advanced search, available as an option, might require some hints to achieve best results, but the ICS standard search is ready to use from the very first moment. As pointed out in the research article by (Lown, Sierra, and Boyer) «a single search box communicates confidence to users that our search tools can meet their information needs from a single point of entry».



Figure 6: Features of the ICS Standard Interface

The ICS Standard interface offers a *Main Search Box* for entering search terms; a *Tool Bar* linking to INIS Home, Thesaurus and Browse; a *Link to Advanced Search*; a *Link to Highlights*, a historical list of announcements made regarding the ICS; and the Language Selector. Both ICS Standard and Advanced Search interfaces are available in eight languages (Arabic, Chinese, English, French, German, Japanese, Russian, and Spanish).

All words put into the query are used; searches are always case insensitive; punctuation is ignored, including @#\$%&^*()=+[] and other special characters; and common articles or determiners, i.e. stop words, such as 'the', 'a', and 'for', are usually ignored. The inclusion of stop words for languages other than English is also possible.

6.3 Advanced Search

The Standard Search interface provides excellent results for most search requirements. Our statistics show that the overwhelming majority of visitors (99.5%) use the ICS Standard Search interface, or some other way, such as access through Google Scholar. Many other studies demonstrate that users are not inclined to become expert searchers (Novotni; Wallace; Valentine). However, if a more precise search is needed, a query can be constructed using the query builder form on the Advanced Search page. Query builder also generates a query syntax which appears in the Advanced Search query box. Alternatively, very experienced users can type the query directly into the Advanced Search query box. In this case, the query builder form becomes disabled.

The Advanced ICS Search (Figure 7), offers the possibility to search all words or an exact phrase; to select (include or exclude) specific metadata fields; and to select the language of publications to be covered by the search. It also supports 'range queries' which enable the user to search for results

where field values are between the lower and upper range specified by the query. Range queries are specified using the Range operator, written as two consecutive dots (e.g. year:2007..2009).

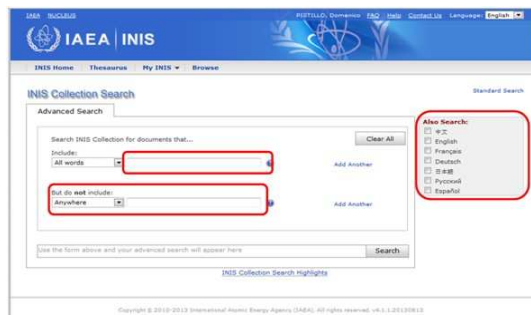


Figure 7: Features of the ICS Advanced Interface

The dropdown menu of the Advanced Search offers the list of fields (metadata elements) in which one can search. This includes Author, Abstract, Title, Country/Organization, Descriptors, Publication year, etc.



Figure 8: Dropdown menu in the ICS Advanced Interface

6.4 Faceted Search

The Faceted Search (Figure 9), or dynamic navigation feature, allows further query filtering by specific metadata like Country, Language, Publication Year, and INIS Volume. Once a search is performed, the dynamic navigation filter options appear in the search results page separated by category. By clicking on a filter option under a category, the search results get filtered to the selected option.

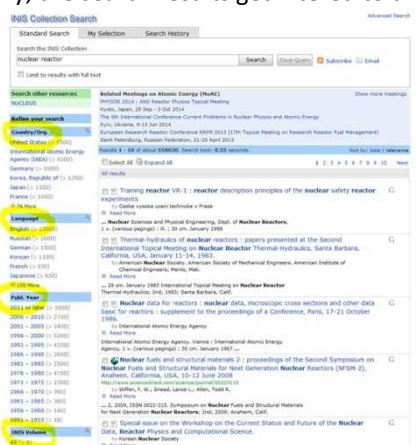


Figure 9: Faceted Search Options

6.5 Expandability

One of the major benefits of using GSA for the INIS Search is its expandability. At the very beginning, ICS searched only the INIS collection of bibliographic records and full texts. However, it was soon realized that the inclusion of the IAEA Library Catalogue, with its 90,000 records, would be beneficial to our users, as well. After the integration of the bibliographic records from the IAEA Library catalogue into the INIS Collection Search, information from the IAEA MoAE database, maintained by the IAEA Nuclear Knowledge Management Section, was also added. The most recent addition was the NUCLEUS database, which provides access to over 130 IAEA scientific, technical and regulatory resources. This

includes databases, document repositories, websites, applications, publications, safety standards, training materials, and more.

Current ICS implementation includes full text PDF documents, bibliographic records, and INIS and PDF metadata. However, the types of documents, as well as their numbers can be expanded. The most obvious expansion is the inclusion of various audio-visual formats, PowerPoint, Word and Excel files, which are not part of the presently deployed application.

6.6 Multilingualism

The ICS interface is available in eight languages — six official United Nations languages (Arabic, Chinese, French, English, Russian, and Spanish), plus German and Japanese. It is also integrated with the INIS/ETDE Multilingual Thesaurus (the same eight languages), which allows searching in different languages using the terms entered. Translation of bibliographic records into other languages is enabled through the integrated Google Translator. Although the translation is, perhaps, not the most satisfying, it gives users an idea about the content of the records found.

6.7 Authority Files

Authority control of a cataloguing database is essential to support the effective creation of adequate metadata (Mainconico), but it is also important for easier, better and faster searches. Keeping that in mind, and with the full integration with the INIS/ETDE Thesaurus, the ICS has fully incorporated an additional eight authority files. They are:

-Journal Title	-Descriptor (DEC, DEI in all languages)
-Journal CODEN	-Country/Organization (of publications)
-Journal ISSN	-Author
-Subject Category Code	-Report Number

Figure 10: Faceted Search Options

6.8 Usability

Usability improvements and the addition of new functionalities increased with the introduction of the GSA-based ICS. Users can now print and export results in different formats, such as PDF, HTML, Excel, and XML, which was not previously possible. At the same time, they can also select individual records and even the fields (metadata elements) to be exported. Another usability improvement, especially for researchers and those working on preparing articles for publishing, was the option to download citations in plain text, RIS format, RefWorks, or as an EndNote. Creation of RSS feeds and the option to email selected search results as a link were also implemented.

6.9 User Profiling

The single most important and popular option is user registration and the creation of user profiles. This was done through a single sign on feature which controls access not only to ICS, but also to a number of other information resources and repositories available through NUCLEUS. In other words, once the user is registered, his registration can be used across a number of resources offering easy single access and further personalization of specific options for each resource. ICS uses personalization to select and remember the interface language, the number of displayed results per page, and to save queries, search updates, and email query results. Connected to user profiling is also a “Workspace concept” where documents which are found are associated with the user profile, and possibly translated into other languages.

6.10 Help

A complex system of aid tools was created and made available on the ICS, mainly through a Help function. It includes FAQ on INIS and a separate FAQ on ICS. The on-line help file covers two dozen elaborated topics, each geared towards a specific ICS feature or option. There are also pop-up hints — examples of how to build a query using metadata, and a recently developed ICS e-training course. Besides these aid tools, a hyperlink leads users to the INIS Collection Search Highlights, which shows a complete set of web announcements regarding the ICS from its inception in 2011. In a way, these announcements represent a brief history of the life and developments related to the ICS.

7 GSA Advantages and Disadvantages

Major advantages of using GSA are users' familiarity with a Google-type interface; the possibility to include many features in the foreground or background; its scalability; quick and relevant response to searches; many out of the box features that can be deployed with minor configuration and easy customization of the user interface by editing its XSLT. In addition to these important advantages, there are also some disadvantages which need to be mentioned. The main one is cost. Google charges according to the number of records (price per each million records) for a GSA license.

In addition, there is the cost of deployment, customization, and daily running and maintenance. It should also be mentioned that the GSA index and/or related database is not in the control of the system administrator since there is no direct access to it, nor is there any access to the search algorithm. One of the most interesting features of Google search is records relevance ranking¹⁸, which many believe is based on hyper-links, among other criteria, and is not easily done within a closed collection, such as the INIS Collection. However, the relevance sorting is used for the display of ICS records, but it remains in a way a "black box". One of the disadvantages often mentioned, and easily noticed by more experienced library catalogue users, is a GSA limitation in building queries, namely, use of a wild card (*) search, which is not supported. Another way to look at the GSA advantages and disadvantages would be to compare some features of the next generation library catalogues, as listed by (Yang and Hofmann)¹⁹

As seen in Figure 11, most of the features are covered in the INIS Collection Search, and even the ones which are not currently included, can be supported by GSA. However, some of the additions would require developmental work on the interface and background system logistics.

Feature	Status
Single point of entry for all library information	Yes
State-of-the-art web presence	Yes
Enriched content	No
Faceted navigation	Yes
Keyword search	Yes
Relevancy	Yes
Did you mean...?	Yes
Recommended/related materials	No
User contribution	No
RSS feed	Yes

Figure 11: Features of the next generation catalogue

8 Conclusions

The decision to choose GSA as a replacement for the previous INIS classic library on-line catalogue was not an easy one. A number of available search engines were examined and evaluated, a team of various specialists was involved, and it was decided to select GSA. Although it might have seemed risky at the time, the last three years of GSA use has shown that it was the right decision that has brought considerable benefits, primarily to INIS Collection users, as well as to the INIS Secretariat — the managing body of the document collection and the search system. User satisfaction was obvious from the results of the survey conducted and from the comments received. The number of visits dramatically increased and most importantly, the number of full-text document downloads also went up. The initial increase in users came with the introduction of GSA, and was further increased by connecting INIS to WorldWideScience.org, and, finally, by making the INIS Collection available on Google Scholar.²⁰

The number of searches increased five times and the number of downloads increased more than ten times²¹. Finally, it should be emphasized that GSA is a search tool. It is not a library collection management tool, not a reporting tool, and not a statistical tool.

However, it works perfectly with Google Analytics for generating metrics and statistical reports. In conclusion, the case of the INIS Collection Search shows that using GSA to search digital library collections is a good choice to attract more users and increase the use of available information resources and repositories.

References

- Chambers, Sally, ed. *Catalogue 2.0: The future of the library catalogue*. 2013. (Cit. on p. 80).
- IAEA. "The Statute of the IAEA". <http://iaea.org/About/statute.html>
- INIS. "INIS Progress and Activity Report". <http://www.iaea.org/inis/about-us/activities/2012-Activity-Report.pdf>
- Lown, Corry, Tito Sierra, and Josh Boyer. "How Users Search the Library from a Single Search Box". *College & Research Libraries* 3. (2013): 227–241. (Cit. on p. 72). Mainconico, S. Michael. "Bibliographic Data Base Organization and Authority File Control". *Wilson Library Bulletin* 5. (1979): 36–45. (Cit. on p. 77).
- Mercun, Tanja and Maja Žumer. "New Generation of Catalogues for the New Generation of Users: A Comparison of Six Library Catalogues". *Electronic Library & Information Systems* 3. (2008): 243–261. (Cit. on p. 80).
- Novotni, Eric. "I Don't Think I Click: A Protocol Analysis Study of Use of a Library Online Catalog in the Internet Age". *College & Research Libraries* 6. (2004): 525–537. (Cit. on p. 73).
- Tramullas, Jesus and Piedad Garrido, eds. *A Review of "Library Automation and OPAC 2.0: Information Access and Services in the 2.0 Landscape"*. 2013. (Cit. on p. 80). Valentine, Barbara. "Undergraduate Research Behavior: Using Focus Groups to Generate Theory". *Journal of Academic Librarianship* 5. (1993): 300–304. (Cit. on p. 73). Wallace, Patricia. "How Do Patrons Search the Online Catalog When No One's Looking?" *RQ33* 2. (1993): 239–252. (Cit. on p. 73).
- Yang, Sharon Q and Melissa A Hofmann. "The next generation library catalog: A comparative study of the OPACs of Koha, Evergreen, and Voyager". *Information Technology and Libraries*. (2013). (Cit. on p. 80).

¹ Google Annual Search Statistics. <http://www.statisticbrain.com/google-searches>

² <http://www.loc.gov/ils/>

³ <http://www.iaea.org/About/about-iaea.html>

⁴ <http://www.iaea.org/Publications/Reports/Anrep2012/orgchart.pdf>.

⁵ <http://www.iaea.org/inis/about-us/membership.html>.

⁶ <http://inis.iaea.org/search>

⁷ <http://nkp.iaea.org/INISSubjectCategories>

⁸ http://en.wikipedia.org/wiki/Basis_database

⁹ <http://www.google.com/enterprise/search/products/gsa.html>.

¹⁰ <http://inis.iaea.org/search>

¹¹ <http://www.iaea.org/inis>

¹² <https://itunes.apple.com/us/app/ne-news/id682405288?mt=8>

¹³ <https://itunes.apple.com/us/app/ne-news-for-iphone/id722935995?l=it&ls=1&mt=8>

¹⁴ <https://play.google.com/store/apps/details?id=org.iaea.nenews>

¹⁵ <http://jolifukyu.tokai-sc.jaea.go.jp/ird/english/index.html>

¹⁶ http://www.iaea.org/OurWork/ST/NE/NEFW/Technical_Areas/RRS/home.html

¹⁷ <http://www.iaea.org/NuclearPower/FR/index.html>

¹⁸ To influence rankings, GSA uses a result biasing policy. A result biasing policy determines the source biasing, date biasing, and metadata biasing settings that are used with a front end. A default result biasing policy is built into the search appliance. Administrator can use default policy, or create one or more custom result biasing policies. A result biasing policy is specific to a front end, so it can be aimed at specific types of end users

¹⁹ Also referred to as "OPAC 2.0" (Tramullas and Garrido), the "library catalogue 2.0" (Chambers) or "the third generation catalogue" (Mercun and Žumer).

²⁰ Impact of joining Google Scholar: 2013: 50,000 searches and 3,000 downloads per month; Currently there are 270,00 searches and 30,000 downloads per month.

²¹ <https://twitter.com/INISsecretariat>

Grey Literature in European Commission Projects^{*}

Sara Goggi, Gabriella Pardelli, Silvia Giannini and Stefania Biagioni (Italy)

Abstract

The survey is focused on the documentation produced by the European Commission (EC) projects involved in the Framework Programme for Research and Technological Development (hereafter FP7) and managed by the Italian National Research Council (hereafter CNR). In particular, the Grey Literature (GL) available on CORDISⁱ and European Projects websites was analysed. In order to verify how it is managed and whether it is compliant with EC recommendations, some categories were introduced to identify, measure and evaluate the usability and availability of projects production. Data was obtained from a sample of European projects websites.

1. Introduction

The latest recommendations issued by the European Commission go towards the revision of their policy on dissemination and preservation of scientific information in order to promote the access to the results of the community-funded research by especially implementing the open access policy within 'Horizon 2020', the EC Framework Programme for Research and Innovation, 2014-2020 (European Commission, 2012). In the proposal for a Regulation of the European Parliament and of the Council laying down the rules for the participation and dissemination in 'Horizon 2020 we can read:

"The rules regarding intellectual property, exploitation and dissemination have been modelled on the widely acknowledged Seventh Framework Programme provisions with further improvements and clarifications. Specific new emphasis has been put on open access to research publications and an opening was made for experiments with open access to other results. The enlarged scope and new forms of funding as well as the need for flexibility in this area of the rules has been taken into account by the possibility to lay down additional or specific provisions where appropriate. Access rights for the European Union, and in the field of security research also for Member States, have been foreseen" (European Commission, COM 2011).

The websites of the EC projects represent an essential vehicle for both the acquisition and the diffusion of documentation and could also become an important resource within a European infrastructure able to overcome the disconnected and scattered nature of their content in order to optimize their reuse. The EC considers the structure and the contents of the websites generally *[suffering] from a contextual and structural neglect* (European Commission, *Best practices*).

In order to improve the documentation management, EC published some guidelines that proved to be a very useful tool for optimizing and handling information on the dedicated portals of the community-funded projects.

The guidelines consist of general and specific recommendations for the websites and for each element of the websites structure. The general recommendations focus on the importance of using social media as well as webmaster tools and virtual meeting facilities (as web streaming) and of adopting a "eu" domain. Among the general recommendations are also identified: the use of sustainable systems aimed at the retrieval of sites by the search engines, the update of the sites contents and the use of keywords in metadata and contents page. We concentrated on the specific recommendations providing the "Best Practice" for the structure of websites, identifying the components and providing the explanation of each concept and content.

The CNR has an important role in the development of the European Union programme: it considers the programmes of the European Union as a priority and strongly encourages the participation of researchers. The CNR has participated in European projects since the beginning and since 2007 joined the FP7 (2007-2013). The CNR also *"welcomes the architecture designed by the European Commission for Horizon 2020 (H2020) and considers its structure innovative and suitable to promote European science, technology and innovation"* (CNR, Position on Horizon2020).

Starting from this context, the aim of this survey is to obtain information about the current practices carried out by the partners of the European projects as a means to improve the projects visibility, the usability and the access to their production.

^{*} First published in the GL15 Conference Proceedings, March 2014.

2. Materials and Methods

We measured and estimated the projects documentation and its level of compliance with EC recommendations. CORDIS was the starting point of our exploration and its contents have allowed us to create a sort of “identity card” for each project. In order to get a specific range of data, we selected a subset of European projects from the CORDIS website using the following query criteria:

i) Search for *>Consiglio Nazionale delle Ricerche*

ii) Refinements *>Programme: FP7 >Coordinator Country>Italy >Participant Country>Italy*.

We got 226 projects matching the search criteria: this group of projects was the core of the analysis.

Data was collected between June and August 2013.

The information was extracted from

- ✓ CORDIS directory contents
- ✓ Project site structure and contents

We gathered the

- ✓ Projects production

and analysed the documents by

- ✓ a) Type of products
- ✓ b) Format
- ✓ c) Availability/access

The objective of detecting all types of documents has been achieved through different steps: first we extracted the information provided by the CORDIS directory: Acronym, Title, Grant, Start/End, Status, CORDIS URL, CORDIS keywords, CORDIS objective, CORDIS related links, Total cost, Area, Contract type and verified the existence of the project websites. This checking allowed us to understand immediately how many projects have their own website even if CORDIS does not always provide the link nor precise information on project sites. In this case we made specific searches via Google to verify the site existence.

Secondly, we checked the projects websites in order to gather information about the websites structure and the projects production.

The third step was the categorization of the documents and the defining of the dataset. Finally, there has been the analysis of the results.

3. Explorative phase: CORDIS website analysis

The following tables and figures show some results: the CNR is the Coordinator in 62% of the cases and in 38% is participant.

Projects subset			Projects subset		
Status	N.	%	Projects websites	N.	%
<i>accepted</i>	14	6,2	<i>yes</i>	140	62
<i>completed</i>	99	43,8	<i>no</i>	86	38
<i>execution</i>	113	50	Projects websites by status		
			<i>accepted</i>	1	0,8
Total	226		<i>completed</i>	60	42,8
			<i>execution</i>	79	56,4

Tab. 1. – Projects websites

Compared to the total, half of the projects appear to be in execution while 43.8% has been completed. A small group of projects is only accepted. About 62% of the projects have their own website: the category of accepted projects shows the presence of only one website compared to 13 which have not developed it yet. Some projects have not created a proper website and are using only a few spaces within the sites of the partners institutions. We included in the group of projects with a website only those which make available some information and/or documentation. This type of website affects the total number of projects for the 8,5%. With respect to this sample, the presence of the projects website does not seem closely related to the project in execution.

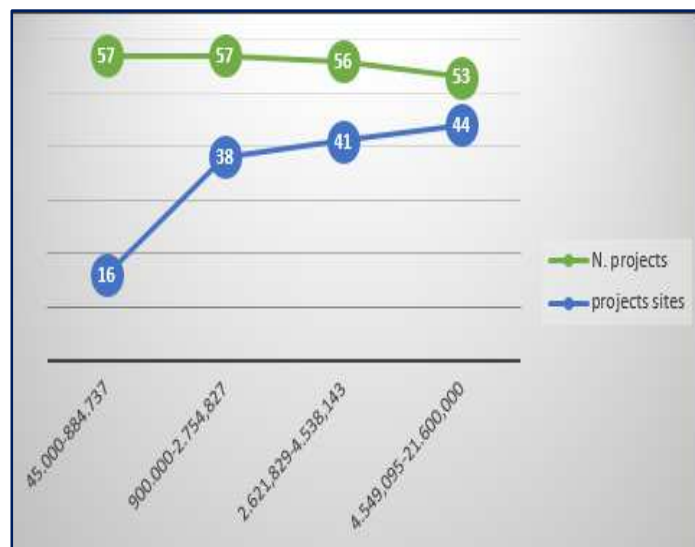


Fig. 1 – Projects websites/projects total cost ranges

It is quite understandable that among the projects recently activated only one has already its own website, but it is not so clear why the other projects with websites are only 56% of the total.

The presence of the project website seems related to the increasing financing; we checked the total cost of projects: when the cost increases, the presence of projects websites increases too (Fig. 1). Probably, the growing financial availability lets to devote more professional management of the sites (e.g.: webmasters, librarians and so on) and a greater number of resources moves towards the Web. Even a higher number of partners requires large and more organized structures and more funding that can help the implementation and the improvement of the project in the network.

The analysis of products managed in CORDIS indicates only two types: publications and reports. CORDIS contains documentation even in the absence of the project website remaining, therefore, the only source of documentation for 24 projects.

There is a large group of projects that does not have any kind of documentation. Then we have the highest percentage of research articles published in journals (Fig. 2), which CORDIS draws directly from OpenAIRE (Open Access Infrastructure for Research in Europe).

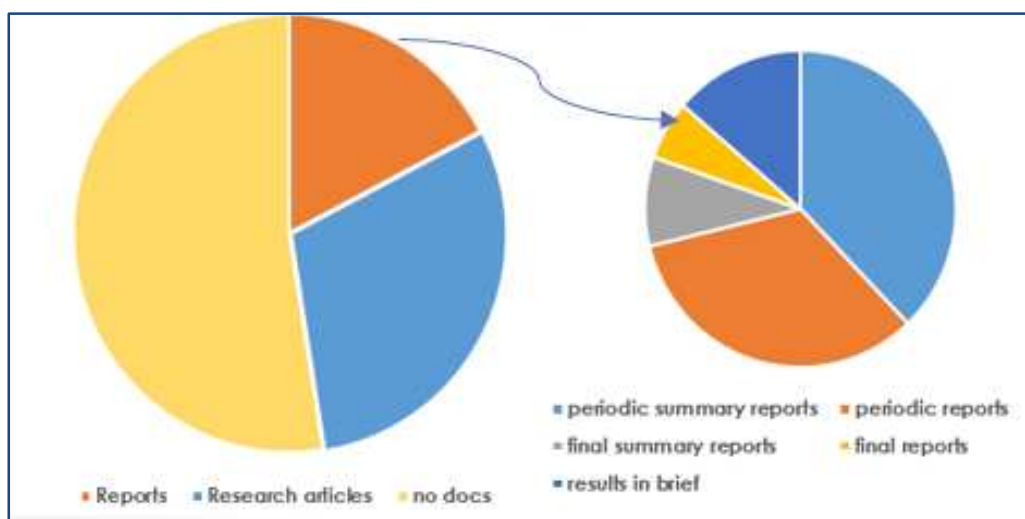


Fig. 2 – CORDIS documentation and reports types

4. Results: projects websites analysis

4.1 Categories

Our next step was the analysis of the projects websites. The collection of contents allowed us to identify the types of documents produced and to gather them into categories, types and subtypes. The nine sections proposed by the guidelines with respect to the structure of the website (Overview - Consortium - Management structure - Case Study, Deliverables, Publications, Events, Media center, Glossary) and the logical association of related materials were the two main criteria followed in the material categorization.

We identified 13 categories, composed by different types and subtypes (Tab. 2).

We treated 7 of them as individual categories, because they can be associated into a single type (Tab. 3).

Categories	n. Projects	%
<i>Courses material</i>	12	8,5
<i>Deliverables & Reports</i>	74	52,8
<i>eNewsletters</i>	31	22,1
<i>Image material</i>	48	34,2
<i>Informative material</i>	138	98,5
<i>Media press</i>	17	12,1
<i>Projects meetings material</i>	50	35,7
<i>Normative documents</i>	2	1,4
<i>Patents</i>	1	0,7
<i>Research/scientific articles</i>	109	77,9
<i>Technical documents</i>	96	68,5
<i>Tertiary documents</i>	20	14,2
<i>Theses</i>	4	2,8

Tab. 2 – Categories

Some categories represent types of materials not specified in the guidelines. We split or merged some categories in order to give more emphasis to the informative content (for example we split *Media centre* in *eNewsletters*, *Image material*, *Media press*, *Normative documents* but merged *Case studies* in *Technical documents* and *Events* in *Informative material*). For the same reason, we preferred to provide a generic level of information about the type *Events*, because their presentation is varied and confusing. Events are almost never divided between past, present or future events and only few of them shows a calendar, as suggested by the guidelines. Very often we found together generic news, not necessarily related to events. Our choice was the creation of the category *Projects Meetings material* whereby we isolated the information concerning the meetings of the projects (workshops, seminars or conferences) thus dividing them from the disorder of the type Event.

Categories	n. Projects	%
<i>Courses material</i>	12	8,5
<i>eNewsletter</i>	31	22,1
<i>Image material</i>	48	34,2
<i>Projects meetings material</i>	50	35,7
<i>Media press</i>	17	12,1
<i>Normative documents</i>	2	1,4
<i>Patent</i>	1	0,7
<i>Theses</i>	4	2,8

Tab. 3 – Individual categories

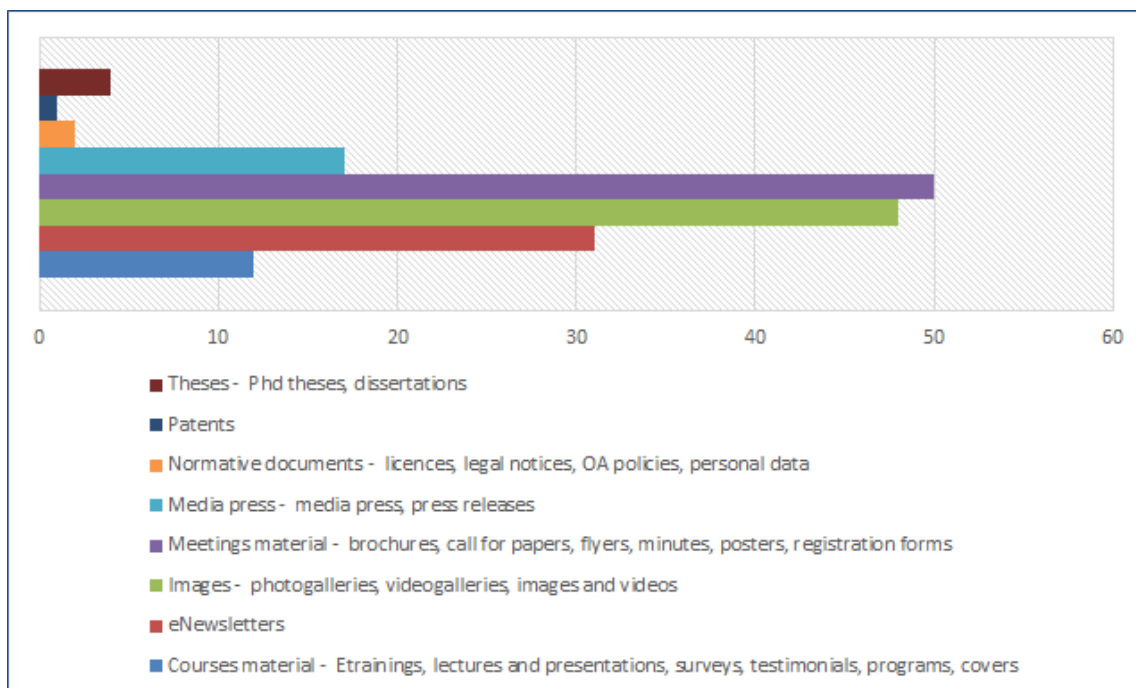


Fig. 3 – Individual categories (subtypes)

Figure 3 shows the types and the subtypes composing the individual categories. *Media press* could flow in the category Information material but it was isolated as contains both newspaper articles devoted to the project as well as articles published by the project itself on national or international press. We considered more interesting to provide this category as separate information. Twelve projects have established courses, training and / or summer school and have chosen to disseminate and make available the related documentation. We checked also several projects that present *Theses* conceived and realized within the project itself. The category *Image material* collects all those projects that have chosen to publish on their own sites various kinds of videos or images not closely related with a specific document type. In these cases, the object itself has been treated as a type. Conversely, when the object is placed as a medium used in support of a given type, we considered it as *pdf*, *html* and *other* formats. The category *eNewsletter* lists the projects that publish their own newsletter in electronic format and make available the full text. Two projects present their normative documents such as licenses, policies and policies for data processing.

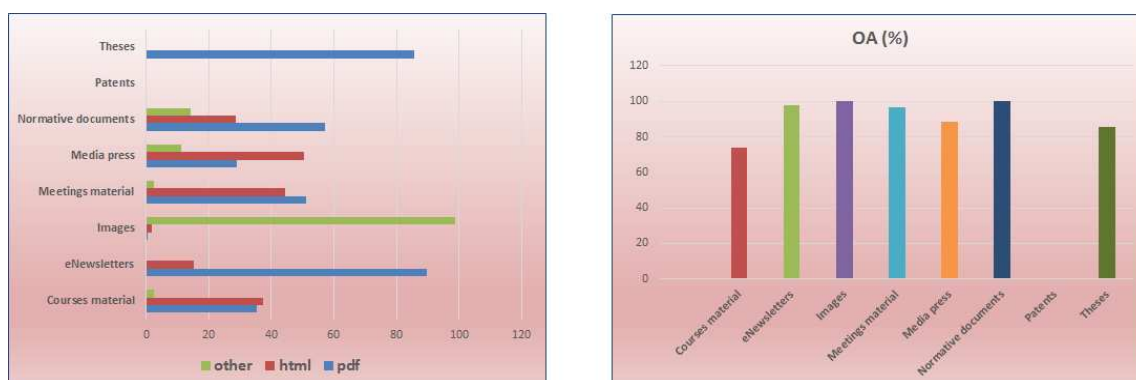


Fig. 4 – Individual categories (formats/accessibility)

The prevalent formats are *html* and *pdf*. The exception is represented the category *Images* which, by its nature, is almost 100% composed of multimedia material fully available. In this case, the formats are jpeg, tiff, png and other video formats. All formats differing from *html* and *pdf* are gathered into the item *other* (Fig. 4).

Talking about access to the full text, we can say that several types make it 100% available (or anyway at very high percentage), so almost all of these types can be open accessed. We could not find any news about the only project that shows a patent because it is only announced but not described, nor viewable.

4.2 Informative material

After the Homepage, the first sections of the guidelines are Project overview and Consortium. Both areas have been included in the category *Informative material*. We can see that almost all projects with a website provide at least a minimum of descriptive information. The contents of these sections are the biggest of the whole category. In Fig. 5 we can see, in fact, the different types and their incidence in the general category. Here we have also materials not mentioned in the guidelines such as Announcements, News, Related links etc.

The objectives of the project are indicated by the guidelines as the basic information for the website design because the treatment of this type of information is highly recommended. They are present almost always although not always introduced by this term. Frequently the objectives are introduced by terms such as: about, overview, abstract, description, mission, but the contents express the objectives of the project.

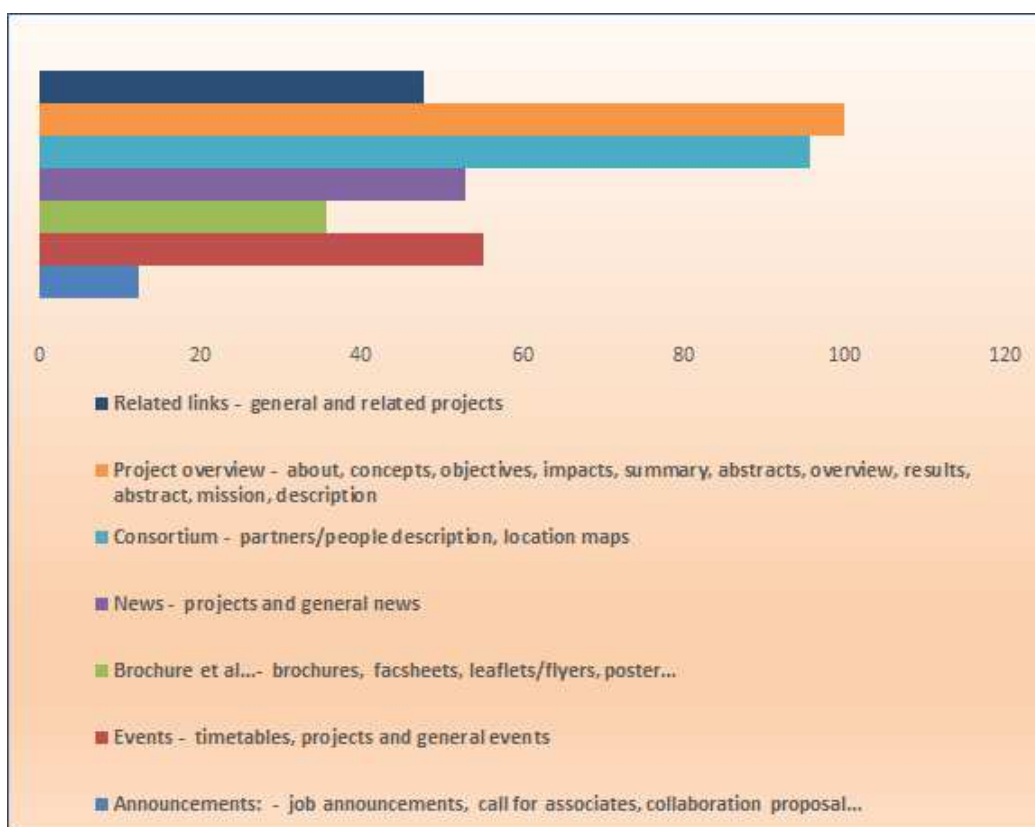


Fig. 5 – Informative material (types/subtypes)

The analysis shows that all projects provide their own overview and that only two out of 140 projects do not give any kind of information on the Consortium.

However, the extent data varies from project to project: most of them propose very little descriptive information while others produce a lot of documentation on scope, objectives, and information on the operational context. The analysis of the completeness of descriptive information shows that about 68% of the projects does not seem to provide detailed data about their operating context. This documentation is almost never provided in a downloadable pdf format, or conforms to a specific pattern. It is given freely and usable as a *html* page (Fig. 6).

Also the presentation of the Consortium varies from project to project. The 95% of the projects indicates the number of participants and their nationality while a slightly lower percentage (87%) is for the contact information that is omitted in some cases.

Significantly less is the presence of location maps, that is the geographic distribution of participants in graphical form. Three projects give a specific presentation about women involved in the project, accompanied by their profile and the description of their activities. In general this is introduced by the term *Gender aspects*.

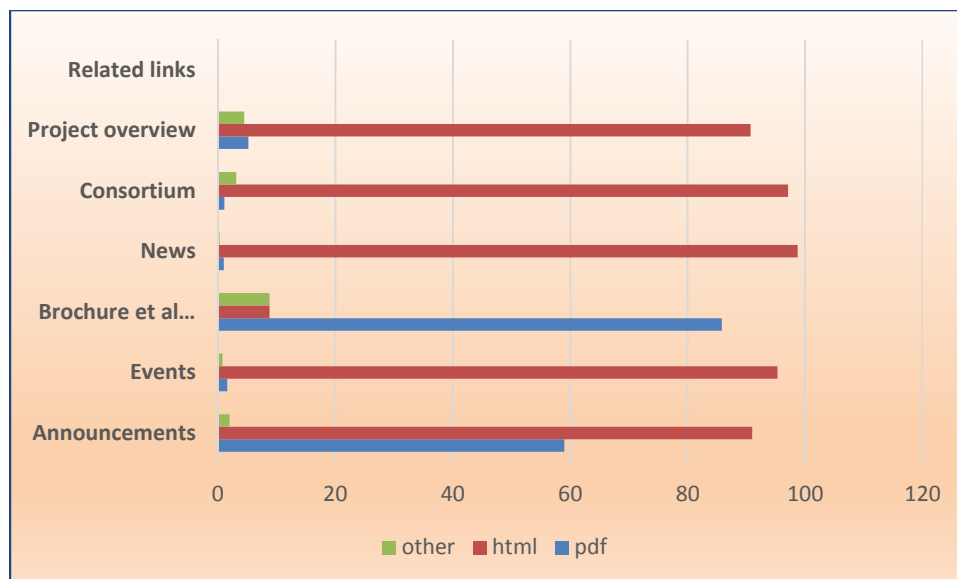


Fig. 6 – Informative material (formats)

Most of the descriptive pages about the partners present the name of the institution, *the* location, the logo and the URL, leaving to the website of the single participating institution the task of presenting the single partners. Several projects, however, perform the description of partners on their own. In this case, the projects propose a detailed description of the activities carried out by the partners and their institution, the subject area covered, the mission and their expertise.

The type Related links indicates the percentage of projects with links to related projects. The format has not been quantified because we have only links to other websites, which were not subject of detection. Data is interesting in order to see how many projects offer the link to similar projects from their websites: compared to the total of the projects with a website, they are about 44%.

4.3 Technical documents

The category *Technical documents* includes some types stated in the sections 3.4 and 3.5 of the guidelines and other types that, in our opinion, are conceptually similar to them, such as roadmaps, prototypes or strategy, just to name a few. The types Research plan, Scientific methodology and Management structures are the largest (Fig. 7).

The projects provide mostly descriptive documentation about their management bodies and about the handling of the project. Approximately, 47% of projects also provide a description of the work packages and/or their management structure. Twentythree projects provide the graphical representation of their research plans, requested from EC guidelines. Also in this case the detail of information is quite varied and the lacking descriptions are numerous. The management structure is presented in different forms and even the terminology used by the projects is not completely uniform and not always explanatory as well as the relationships between government bodies, activities and methodology.

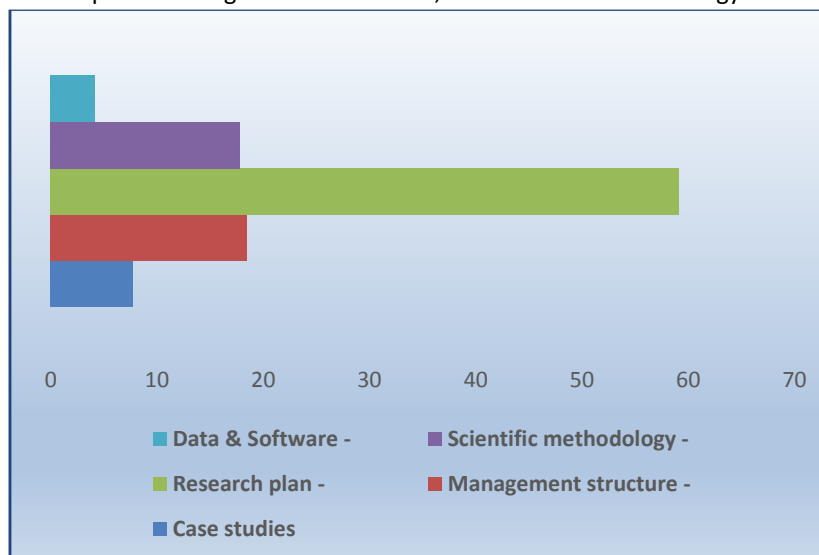


Fig. 7 – Technical documentation (types/subtypes)

A few projects present *Case studies*. It is true that the guidelines recommend the description of *Case Studies* only whether they are particularly relevant. Moreover, in the guidelines, the concept of *Case studies* appears closely related to geographic studies. It is also expressly recommended that the description of *Case studies* should be accompanied by a geographical representation of the same, by the creation of global or European maps. If we give a quick look at the research themes of the projects presenting *Case studies*, we see that mostly belong to the Space and Environment areas. Comparing the total number of projects with a website and the projects that describe *Case studies*, the percentage is very low and their impact within the category *Technical documents* is around 13%. By restricting the calculation to the total percentage of the projects contained in these areas, it appears that 65% of the projects have *Case studies* and 10% of the latter accompanies the description with geographic maps. As for the other types, most of the technical documents are made available through descriptive *html* pages. Some projects use simple links to the studied sites, or to images and photographs of specific subjects. Among *Case studies*, only one project makes available the description of the site.

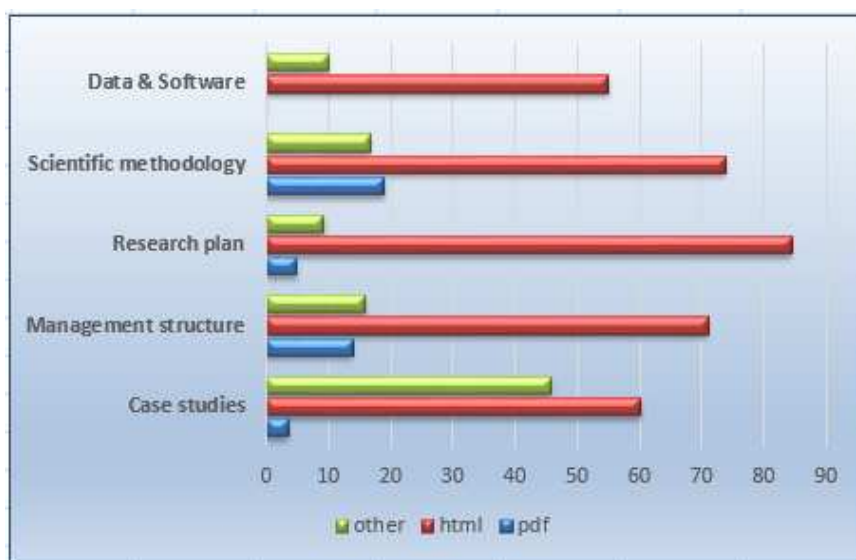


Fig. 8 – Technical documents (formats)

All documents converged in this category are usable in open access.

4.4 Deliverables & Reports

The guidelines give recommendations only about public *Deliverables*. These should be listed in tables and provided with specific descriptive metadata that can be used for the search and retrieval. It is also recommended the use of *pdf* format for downloading. The Research articles should be visible in specific tables (separate from those of the *Deliverables*) identified and described through the common metadata and, if open access, should be downloaded directly from the project website.

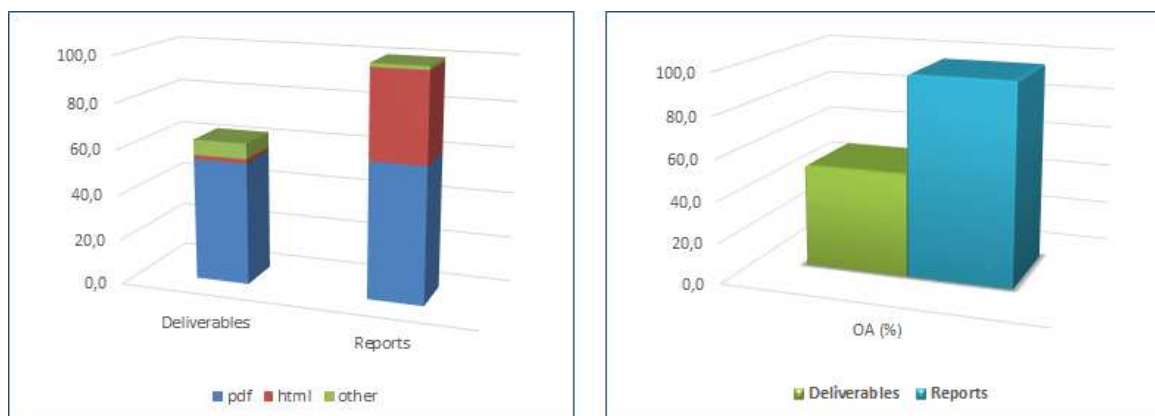


Fig. 9 – Deliverables & Reports (formats/accessibility)

The guidelines do not give indications about *Reports* but almost 60 % of the projects make visible some *Reports* on their sites (Fig. 9).

The subtypes are the same already met in CORDIS: periodic or final reports of activity and results in brief. However, there are also a few working papers and/or technical reports. The *Reports* are freely available in 97% of the cases.

The *Deliverables* are considered the main product of the projects activities. They allow checking the progress of work.

The 54% of the projects present their *Deliverables*. It seems that 46% of projects does not show them but many projects maintain a restricted area so it is possible that all *Deliverables* are deposited inside, without distinction between public and private. It is difficult to understand, because the projects that explicitly state to the partners that *Deliverables* are stored in the reserved area are rare, most of them do not provide any information about the contents of this area. The presentation of the *Deliverables* varies: some projects give news in the form of a list, calling them directly from the page menu; others put them in different voices such as documents or publications; and others give the information together with the work packages.

4.5 Research/scientific articles

The presentation of the *Research/scientific articles* category varies from project to project and is presented using different terms. It is rarely available in tabular form, as suggested by the guidelines, and even more rarely is structured in form of searchable databases. The form of presentation is often a list with links to the publisher's site for the access to the full text. Compared to the *Deliverables*, we noted a bigger compliance to the guidelines about the use of descriptive metadata: generally the bibliography contains the basic elements useful for the resource identification. We verified the there is no full correspondence between research articles appearing on the projects sites and on OpenAIRE: the most frequent case is that in which the project shows a greater number of references than OpenAIRE.

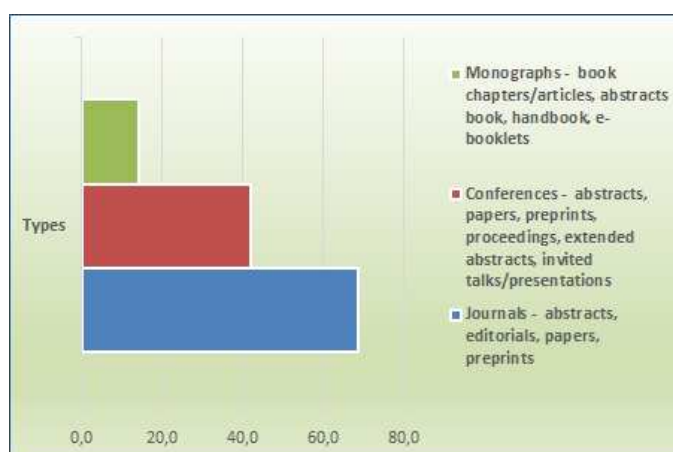


Fig. 10 – Research/scientific articles (types/subtypes)

The highest percentage of the group is related to research articles published in journals, followed by publications in conferences and then in monographs. The subtype shows that the category includes papers, abstracts, editorials and proceedings. OpenAIRE detects articles published in journals. It is rare that OpenAIRE shows articles published in Proceedings.

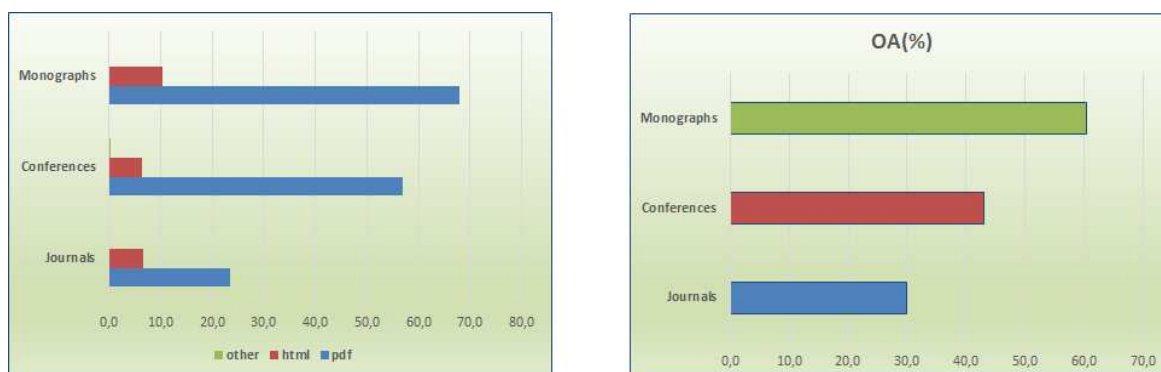


Fig. 11 – Research/scientific articles (formats/accessibility)

The observation of formats and accessibility on figure eleven shows that the more intensive use of the *pdf* and the availability of full-text are attributable to monographs, followed by Conferences and Journals. This is due to the presence of the subtypes e-booklets and handbooks in Monographs and the subtypes abstracts, extended-abstracts and invited talks / presented in Conferences. Most of the books and papers published in journals, book chapters and proceedings are available, but not accessible in full text. The availability of papers in open access is attributable to the presence of the document in Arxiv, in institutional repositories and/or in open access journals.

4.6 Tertiary documents

The category includes all types considered as secondary sources in this context. The projects that show this kind of products are numerically very low. Both the catalogues and the bibliographies are structured in the form of *html* pages (Fig. 12).

The category includes the *Glossaries*, treated in the last section of the guidelines that recommend to projects the production of a glossary of terms and abbreviations and of a Wiki. We checked four projects responding to this recommendation. They make available on their websites the explanation of terms and expressions of their own field of study. The nature of the glossary depends on the scope of the research project. Only one project has structured the glossary as a searchable database.

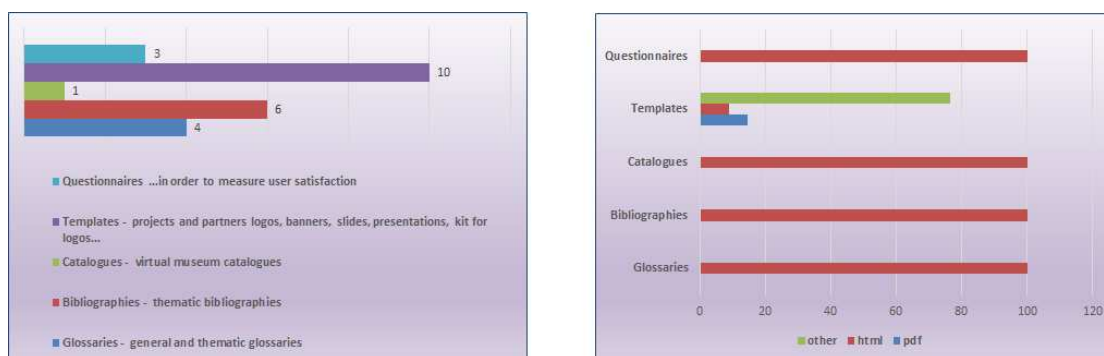


Fig. 12 – Tertiary documents (types/subtypes/formats)

Five projects propose a Wiki, more or less structured and articulated. In general they are organized as all wikies or shared websites and give more advices about the project. Sometimes they create special pages and recommendations about technical details and analysis of the possible approaches to the topics of study.

The questionnaires are offered by three projects with the aim of mapping the intended audience, to check the skills and to acquire information. They also aspire to measure the degree of user satisfaction about the contents of the project and their presentation on the websites.

5. Concluding Remarks

We encountered several difficulties in the identification of the contents offered by the projects. Sometimes the websites are very articulated and the organization of the material is not immediately understandable. In order to provide clear and explanatory results, we collected the documents in some categories. The *Grey Literature Vocabulary* (Pejšová P. et al., 2012), was our starting point and our guide in categorizing: more or less all types of materials encountered in the data collection found their collocation within the categories proposed by the vocabulary.

The analysis revealed that the documents produced by projects and available from their websites are numerous and more varied than the items included in the EC guidelines and that there are important differences in the way the contents are disseminated. Resuming, we can say that there is a basic compliance to EC guidelines in terms of content but a minor compliance towards the use of *pdf* format, compared to the increased use of *html*, and to the observation of specific templates and models in the preparation of contents. Almost all categories are available in open access apart from the deliverables and the research articles published in journals.

Surely, the European Commission is very interested in the divulgation of projects results and shows a remarkable attention to the communication plan and to the involvement of user groups. Our society is increasingly technological and needy of documentary sources visible and shareable. The attention given by the EC to open access formats and to the retrieval of information resources of different domains, seems to be a confirmation of the will that knowledge becomes a common heritage, not only accessible by an elite of specialists. A contribution in this sense could come from the more extensive use of tools such as Forums, Blogs and Social Networks. Our analysis reveals that 14% of the projects has already

created a Forum and/or a Blog and that 17% is included in the circuit of the most common Social Networks such as Facebook, LinkedIn and Twitter.

In our opinion this is a good result. The use of these tools can provide alternative routes and give more visibility to the realities of the projects and the supported activities. The sharing of collaborative platforms may facilitate the activities, the insights and the comparisons and act as a major vehicle of information at all levels.

Another important topic of reflection is related to the mode of content presentation. We detected a wide heterogeneity on the choice of terms which label the documentation: documentation itself, publications, library, results, dissemination are only some of the terms used by the projects to collect and make available their documents. This heterogeneity primarily involves the Grey Literature that is the majority of the documentation produced. Grey Literature remains too often in the shadow. OpenAIRE checks only *Research/scientific articles* produced by the projects, but any type of grey documents is gathered.

The scholars involved in European projects need to be updated on the state of the art of the research conducted by other projects: current, completed and/or belonging to different and past frameworks. They need of all types of documents. So, it becomes very important the creation of infrastructures that merge, describe and make easily available all documentation, grey literature included.

Equally helpful may be the adoption of taxonomic criteria for extracting from the sites informative redundancy. So, the resources could be labelled with more strictness.

Harmonize the terminology of the categories, especially of GL categories within each project, through the use of guidelines careful to the taxonomic criteria could be of great help to the visibility of the documentation.

We detected that also CORDIS, in most of the cases, is the only source for establishing links between projects belonging to the same thematic areas but these areas are very general and do not specify the topics studied by the projects. A significant example in this sense is given by thematic areas such as Scientific Research, Information Systems or Coordination and Cooperation, which do not appear to reach a significant level of detail to guide the search for users. The projects websites do not offer links to related projects by research areas or specific subjects. Only one project publishes on its website the keywords of the studied topics and only three projects present Tag Clouds, that could be used to visually represent the keywords used in a website and the related fields of research projects.

We do not believe it is our task to make further recommendations and/or suggest their extension to improve the visibility and accessibility of the vast documentation produced by the projects. We only provide a map of the results and some thoughts on what emerged from the analysis but we hope that European Commission can work at the creation of infrastructures careful to the relationship between the content itself and accompanied by adequate tools to search and retrieve information. We retain that the use of controlled vocabularies, a greater emphasis on conceptual content and the use of descriptive and semantic metadata that can label and describe the documents in a timely manner could be very important. Structures designed this way could collect and make more accessible the whole documentation.

6. Future plans

In the future, we would like to extend this study to other realities. For example, it could be very interesting to analyse projects coordinated by partners from other countries and/or by different institutions, such as the universities or the industries. The study of the terminology adopted by the EC inside the different frameworks and in the projects websites it could be equally interesting. Moreover, we found that among the information provided by CORDIS, particularly stimulating are those related to the involved research/thematic areas of the projects. The areas cover the majority (approximately 83%) of the total of areas involved in the seventh Framework Programme. Most of the projects pertains to the following areas: *People* - devoted to the "human potential" in Europe; *ICT* - devoted to Information and Communication Technologies; *NMP* - devoted to Nanosciences and Nanotechnologies; *IDEAS -ERC* - devoted to "frontier researches". Among these four areas, *ICT* and *NMP* have an equally significant number of websites. Conversely, in *People* - that is the area of greatest concentration - only 21% of the projects shows to have their own website. For some areas, which count a relatively small number of projects, the number of websites is coincident with one of these projects, for which the hedge is total. Topics and thematic areas in relation 1:1 are the majority, but some topics are shared between different thematic areas.

It was not possible, here, to study them in depth, therefore we just gave a few indications. For this reason, a deeper analysis of the relationships between projects, websites and research/thematic areas could be another possible directions for further development and future studies.

References and Webliographyⁱⁱ

- CORDIS - Community Research and Development Information Service
<http://www.cordis.europa.eu/projects/home_it.html>
- EUROPEAN COMMISSION. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Towards better access to scientific information: Boosting the benefits of public investments in research*. Brussels, 17.7.2012 COM (2012) 401 final.
<<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0401:FIN:EN:PDF>>
- EUROPEAN COMMISSION. *Proposal for a Regulation of the European Parliament and of the Council laying down the rules for the participation and dissemination in 'Horizon 2020 – the Framework Programme for Research and Innovation (2014-2020)'*. Brussels, 30.11.2011 COM(2011) 810 final 2011/0399 (COD).
<<http://www.europarl.europa.eu/document/activities/cont/201203/20120316ATT41112/20120316ATT41112EN.pdf>>
- EUROPEAN COMMISSION, Directorate General, Research & Innovation, Directorate I , Environment. *EU Project Websites – Best Practice Guidelines*.
<http://ec.europa.eu/research/environment/pdf/project_website_guidelines_en.pdf>
- CONSIGLIO NAZIONALE DELLE RICERCHE, *Position of the National Research Council of Italy (CNR) on HORIZON2020*.
<http://www.attivitaeuropee.cnr.it/sites/default/files/antonella/POSITION%20PAPER%20on%20H%202020%2016052012_1.pdf>
- EUROPEAN COMMISSION - Research & Innovation. *Horizon 2020*
<http://ec.europa.eu/research/horizon2020/index_en.cfm?pg=h2020-documents>
- EUROPEAN RESEARCH COUNCIL. *Supporting top researchers from anywhere in the world*.
<<http://erc.europa.eu/media-and-events>>
- PEJŠOVÁ P., MYNARS J., SIMANDLOVÁ T.. *A linked-data vocabulary of grey literature document types: Version 1.0*. Thirteen International Conference on Grey Literature (GL13), Washington D.C. (US), 5-6 December 2011, TextRelease, Amsterdam, 2012, (GL Conference Series, ISSN 1386-2316; no 13).
<<http://code.google.com/p/grey-literature-typology/>>
<http://invenio.nusl.cz/record/81435/files/idr-146_2.pdf>
- POCH M. *PANACEA Platform documentation index*. WP3. Internal report 3.0, 2012.
<http://www.panacea-lr.eu/system/tutorials/PANACEA-Platform_documentation_index_v3.0.pdf>
- SCHÖPFEL J., 2010. *Access to the European Grey Literature*, in PetraPejšová ed., Grey Literature Repository, RadimBačuvčík, VeRBuM,2010.
<<http://nrgl.techlib.cz/images/Book.pdf#page=20>>
<<http://www.panacea-lr.eu/en/info-for-researchers/data-sets/>>
<<http://libguides.gwumc.edu/content.php?pid=44946&sid=4110551>>
<<http://cordis.europa.eu/search/index.cfm?fuseaction=search.simple>>
<ftp://ftp.cordis.europa.eu/pub/fp7/ssh/docs/synopsis_ssh_projects_fp7_low_definition_with_web_links.pdf>
<ftp://ftp.cordis.europa.eu/pub/fp7/ssh/docs/ssh-project-synopsis-fp6_en.pdf>
<<http://cordis.europa.eu/fp6/>>
<http://cordis.europa.eu/fp7/ict/programme/home_en.html>
<http://cordis.europa.eu/search/index.cfm?fuseaction=prog.document&PG_RCN=13628900>
<<http://bibliontology.com/examples>>

ⁱ Community Research and Development Information Service, is an information space devoted to European research and development (R&D) activities and technology transfer. http://cordis.europa.eu/guidance/helpdesk/faq_en.html

ⁱⁱ URL last access: November 21 2013

Effect of Open API, NDSL Open Service (NOS) on Sharing Technical Reports in Korea^{*}

Seon-Hee Lee and Mi Hwan Hyun (Korea)

Abstract

This paper is a case study on sharing technical reports through Open Application Programming Interface (API) in Korea. In 2009, KISTI implemented an OpenAPI called NDSL Open Service (NOS) to increase sharing of national R&D research results among users through major portals, institutes, and universities in Korea. NOS participants can access the technical reports from their website directly instead of having to use the NDSL platform. Technical report usage has increased dramatically since the implementation of NOS. This study analyzes the effect of NOS on the national R&D technical report usage by examining the changes in the amount of usage and the number of participating organizations in NOS. The number of universities, institutes, and portals using NOS and the usage statistics for technical reports through NOS was examined from 2009 to 2012. The results revealed that the number of participating organizations in NOS increased 191% from 23 in 2009 to 67 in 2012. The amount of usage of technical reports through the NOS has increased 83% since its implementation over the previous usages by the NDSL portal alone. Thus, National R&D technical reports are being shared among users through various portals, universities, research institutes, and companies through NOS. NOS has contributed to populating grey literature repositories in Korea. Through NOS, Korean R&D research results as a national knowledge asset can be utilized by a vast number of users for creating new studies and avoiding duplication of research.

Introduction

This paper is a case study on sharing technical reports through the Open Application Programming Interface (API) in Korea. Korea Institute of Science and Technology Information (KISTI) has been collecting the technical reports of national research and development projects through the National Science & Technology Information System (NTIS). KISTI has shared technical reports with the users through a portal system called the National Discovery for Science Leaders (NDSL). NDSL contains about 100 million records of Science and Technology Information (STI). As of Nov. 2013, KISTI has provided about 68 million journal articles, 29 million patents, 170,000 titles of national R&D technical reports, and more. In 2009, KISTI implemented an OpenAPI called NDSL Open Service (NOS) to increase sharing NDSL contents including the national R&D research results among users through major portals, institutes, and universities in Korea. Before 2009, if users wanted to access NDSL contents, they had to visit the NDSL website.



< Figure 1> NOS website (<http://nos.ndsl.kr>)

^{*} First published in the GL15 Conference Proceedings, March 2014.

NOS is an open platform. NOS participating organizations can access, connect, and download the technical reports from their website directly using OpenAPI instead of visiting on the NDSL platform. Technical reports usage has increased dramatically since the implementation of NOS.

Goal

This study analyzes the effect of NOS on the national R&D technical report usage by examining the changes in the amount of usage and the number of participating organizations in NOS.

Method/Process

Configuration of NOS was introduced for its design and system. Data exchange protocols and usages of NOS were compared among types of resources.

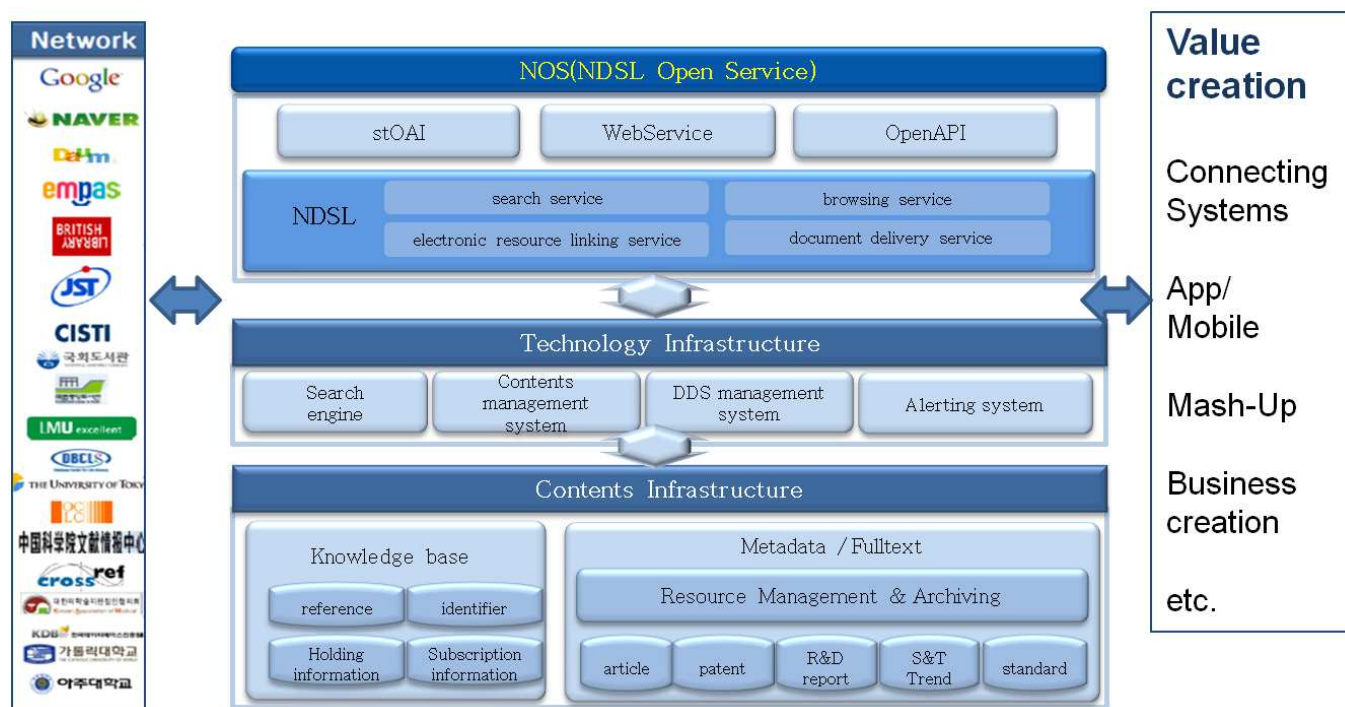
The number of universities, institutes and portals participating in NOS and the technical report usage through NOS was examined from 2009 to 2012.

The usage statistics were obtained from the NOS usage monitoring system.

NDSL and NOS technical report usage were compared.

NOS overview

NOS Configuration



< Figure 2> Configuration of NOS

NOS utilizes contents and services in NDSL infrastructure as seen in the NOS configuration <Figure 2>. NOS provides metadata and the full-text of R&D technical reports, articles, patents, S&T trends, etc. from NDSL Resource Management & Archiving. NOS uses references, journal holding information, identifiers, and subscription information of specific journals from the knowledge base. NOS provides searches, browsing, electronic resource linking services, an alerting system, and document delivery service.

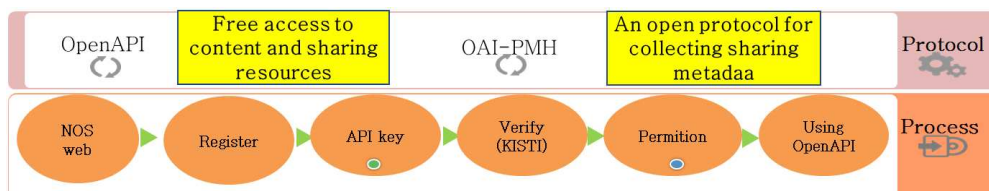
The NOS management system provides the following functions: management of authentication keys, identification of participating organizations, service checking overloading, and usage statistics. Due to increasing usage of NDSL contents by NOS local organizations, usage of NDSL contents has increased as well. Therefore, system scalability should be enhanced for sustainable search results and utilization for NDSL contents by local organizations. To address these needs, KISTI developed a virtual server based on cloud computing as the NOS system infrastructure.

NOS Protocols & Usages

<Table 1> Comparison between stOAI and OpenAPI

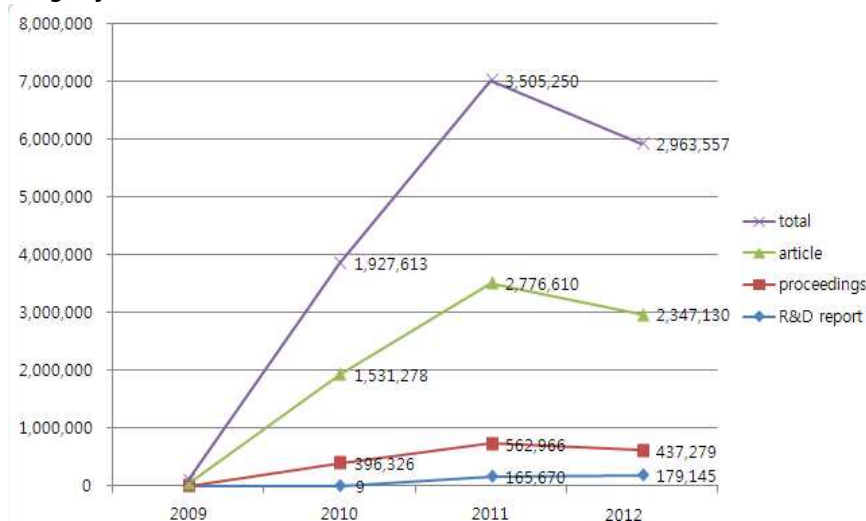
Type	stOAI	OpenAPI
Protocol	OAI-PMH	REST
Contents	domestic articles, R&D reports, S&T trends	articles, patents, R&D reports, S&T trends, standards
Services	metadata, full-text links	search, browse, DDSs(Document Delivery Service), link resolver
Data transmitting method	XML	RSS XML JSON
Authentication method	IP authentication	API Key authentication

NOS is sharing and opening contents of NDSL in two types of standards technologies. stOAI is based on the OAI-PMH protocol and OpenAPI is based on the REST protocol. OpenAPI and stOAI are different in scope and in their use procedures of the content opened as in <Table 1>. In stOAI, institutions are authenticated through IP, and only the institutions that have agreed in advance through MOU can use the service. OpenAPI can be used by any user if an authentication key is issued through the NOS website (<http://nos.ndsl.kr>) without any particular restriction as seen in <Figure 3>.



< Figure 3> OpenAPI and OAI-PMH

Usage of NOS



< Figure 4> Downloads by types of resources through NOS

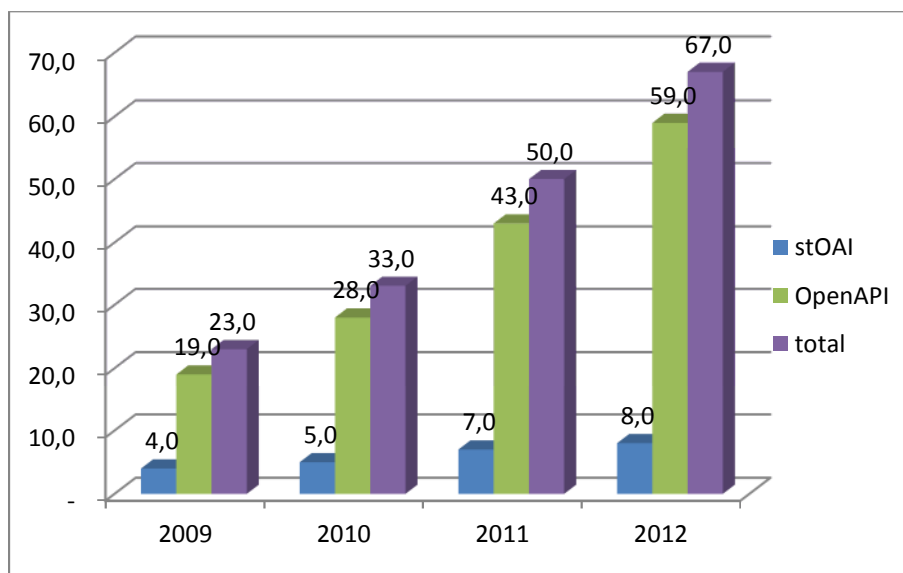
The NOS usage is demonstrated in <Figure 4>. It has dramatically increased since 2009. The total number of NDSL contents downloads was 3,505,250 in 2011 and 2,963,557 in 2012. The most popular download type of content was journal articles. The ratio of journal article downloads was 79.2% in 2012. Therefore,

the patterns of journal article downloads and total downloads are similar. The number of R & D technical report downloads has increased rapidly from 9 records in 2010 to 179,145 records in 2012.

Results

NOS participating organizations

The number of NOS participating organizations has increased 191% from 23 in 2009 to 67 in 2012 as seen in <Figure 5>. The number of NOS participating organizations has increased every year.

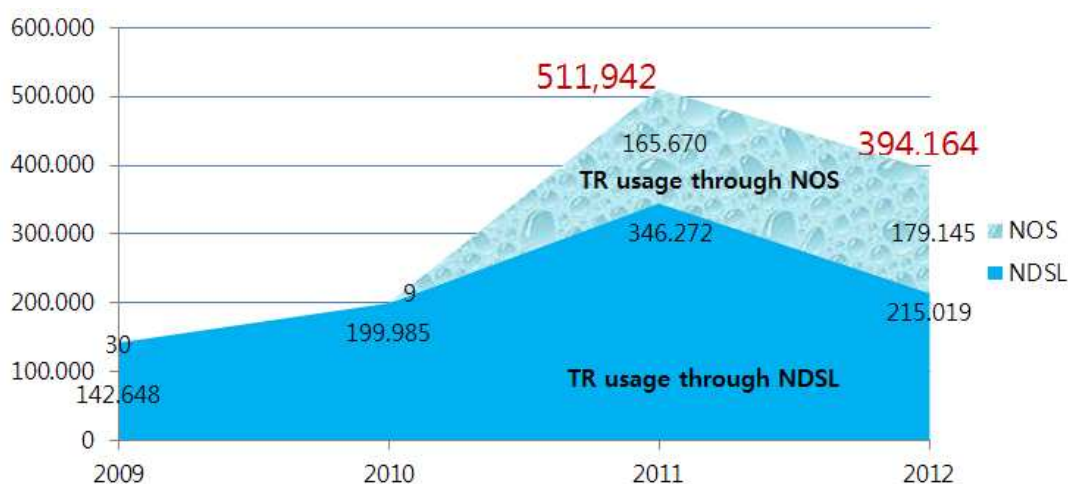


<Figure 5> NOS organizations (2009-2012)

Eight organizations were using stOAI in 2012. The number of participating organizations has increased about 100% over that in 2009. However there are fewer organizations using stOAI (OAI-PMH) than those using OpenAPI. Those organizations usually have their own servers, build DB for NDSL contents, and manage them. The ratio of OpenAPI organizations is 88% of the total in 2012. The organizations using OpenAPI have increased 210% since 2009 as seen in <Figure 5>. The reason OpenAPI is more popular than stOAI is that the OpenAPI organizations do not need to build their own system and DB. If an authentication key has been issued to organizations through the NOS website by KISTI, they can use NDSL contents and services. Their users just access the NDSL platform and use the contents. Therefore more users can access more STI contents with fewer expenses.

TR downloads on NOS & NDSL

The technical reports (TR) usage through NOS in terms of full-text download has increased. The number of TR downloads in 2012 was 394,164 titles. TR downloads through NOS has increased by 83.3% since its implementation over the previous usage by the NDSL portal alone as seen in <Figure 6>.

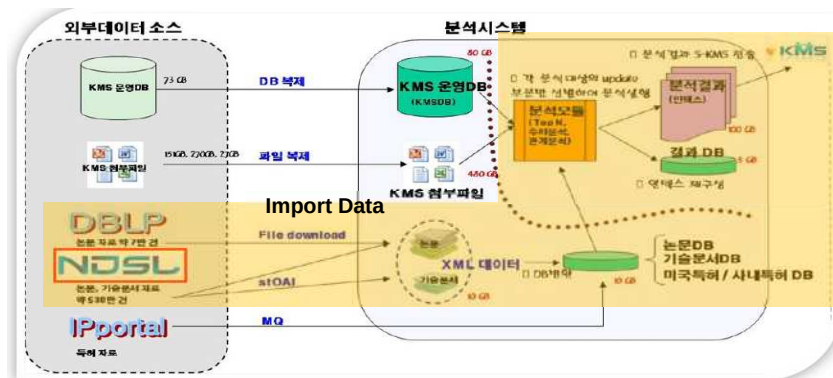


<Figure 6> TR full-text downloads (2009-2012)

Utilizations by Organizations

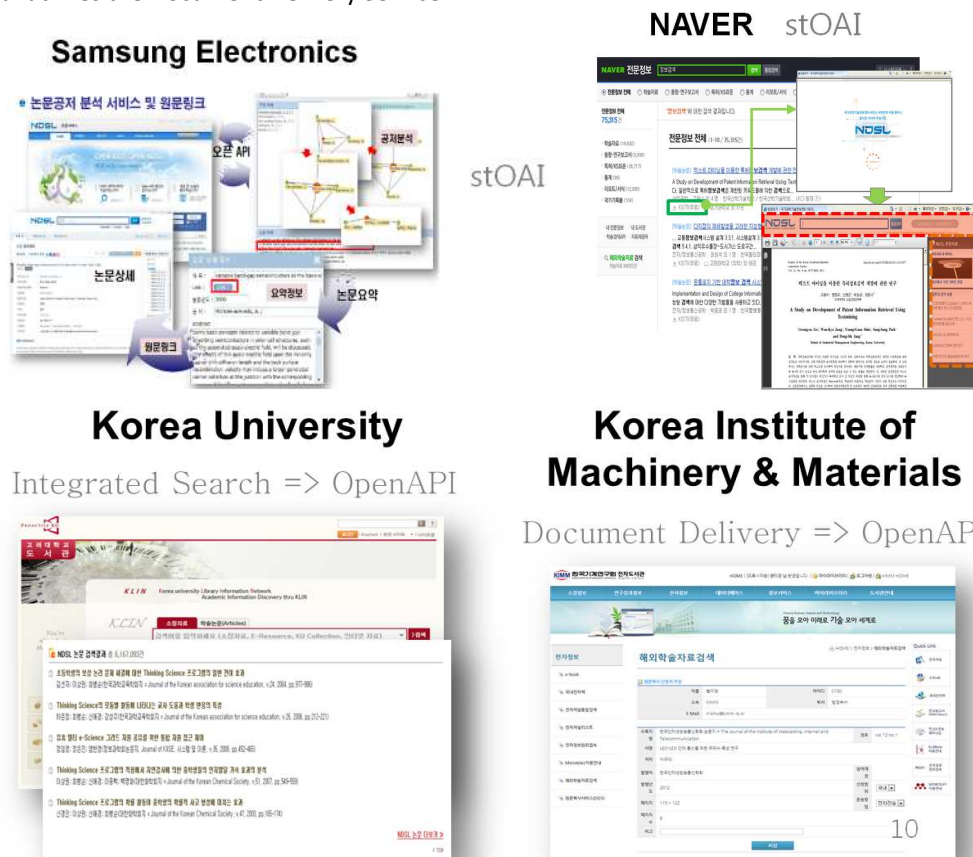
The NOS participating organizations include universities, research institutes, and major portals. Seoul National University, Korea Advanced Institute of Science and Technology, NAVER, Samsung Electronics, and other institutes are NOS members.

Secondary data processing
(keyword extraction)



<Figure 7> Samsung Electronics' Text Mining Processing

Samsung Electronics, Naver, and Exlibris use the stOAI protocol. They upload NDSL contents on their own servers, build DBs, and create services. Samsung Electronics applies text mining to NOS contents to provide better services for users as seen in <Figure 7>. Samsung Electronics provides mash-up services such as technology trend analysis and technology terms relations analysis to their users. Naver, the most popular portal in Korea, provides NDSL contents of metadata and full-text link through NOS. Their users can access NDSL contents without visiting the NDSL website. Exlibris, a global library solution company uploads NDSL contents through NOS on Primo Central, Metalib, and SFX Solution. Exlibris users all over the world from more than 5,000 institutes can use free of charge Korean R&D research results and have linking e-resource service. Institutions upload NDSL contents to their website make their functions stronger. Korea University uses an integrated search service and Korea Institute of Machinery and Material utilizes the Document Delivery Service



<Figure 8> Utilization Cases of NOS

Conclusion

National R&D technical reports produced by Korean government funds are being shared among users through various portals, universities, research institutes, and companies through NOS.

The usage of technical reports has increased dramatically since NOS implementation in 2009 and more than 45% of the total TR usages were made through NOS in 2012.

NOS has contributed to populating grey literature repositories in Korea.

Through NOS, the Korean R&D research results as national knowledge assets can be utilized by a vast number of users for creating new studies and avoiding duplication of research.

Effects of NOS are

- improving R&D capabilities by provision of the latest, qualified STI and value-added services
- promoting distribution of R&D results among organizations such as academia, research institutions, etc.
- budget savings for DB construction and system development, and their management for participating organizations
- providing new mash-up services that are needed for users
- commercialization by participating organizations
- building an ST knowledge ecosystem through creative ideas

References

- [1] Jun-Weon Yoon. 2007. "Harvest and Providing System based on OAI for Science." *Journal of Information Management*, 38(1): 141-160.
- [2] Mi-Hwan Hyun, Soo-Mi Shin, Wan-Jong Kim. 2013. "Study on Sharing and Spreading of S&T information: Focused on the Case of NDSL Open Service." *Korea Convergence Society*, 2(1): 1034-1035.
- [3] Seung-Jun Cha, Kyu-Chul Lee. 2009. "OpenAPI Integrated Management System Development through Service Information Management." *Database Study*, 25(3): 23-40.
- [4] S.L. Shreeves, J. Kaczmarek, T.W. Cole. 2003. "Harvesting Cultural Heritage Metadata Using the OAI Protocol." *Library Hi Tech*, 21(2): 159-169.
- [5] Yong-Ju Shin, Nam-Kuu Kang, Won-Kyun Joo, Ki-Soek Choi, Jae-Soo Kim. 2010. "Open-API Design Study to Link and Use Information of Study Result of National R&D Article." *Korea Society of Internet Information Fall Conference Collection of dissertations*, 11(2): 229-230.
- [6] UIUC (Univ. of Illinois at Urbana Champaign) OAI Metadata Harvesting Project, [cited Nov. 14, 2013] < <http://uilib-oai.sourceforge.net/> >

Effects of legal changes and electronic publishing on the access to GL in Iceland^{*}

Stefanía Júlíusdóttir (Iceland)

Abstract

This article describes briefly legal stipulations on collection building and public access to printed matter (grey literature as well as market publications) in Iceland since a printing press was brought to the country around 1530 (Sigurðsson, 1971). The importance of public access to knowledge and information in a democratic society is argued for. The findings of a study (2002) on the changes in proportions of grey versus white literature from 1944 to 2001 are described. They indicated greatly increase output of written knowledge in Iceland during the latter half of the 20th century, in particular increases in GL, international peer reviewed journal articles, and patent applications. Moreover a suspicion arose that GL might increasingly have been produced in-house, and published electronically as the 20th century drew to a close. If so that would have been a problem because publications produced in-house were often not delivered on legal deposit, due to ignorance of their producers of the law. The e-medium was, furthermore, not covered by the *Legal Deposit Act no. 43/1977* in force until the 1st of January 2003 and therefore access, both bibliographic and to the publications themselves, was difficult or even missing. The findings of the follow-up study in 2012 indicated that contrary to expectations technological developments that facilitated in-house production of GL and e-publishing towards the end of the 20th century and at the beginning of the 21st century had made GL harder to access. Amendments to the *Legal Deposit Act no. 20/2002* were necessary to maintain the kind of public access stipulated in previous legal acts in Iceland and enjoyed by the Icelandic public during the period of print on paper.

Introduction

A printing press was brought to Iceland around 1530 by the bishop at Hólar in the diocese of northern Iceland. Around that time Iceland came under Danish rule, along with Norway under which power it had been since 1262. Although the first act on legal deposit in the kingdom of Denmark only came in 1697, the Royal Library in Copenhagen received a copy of all which was printed in Iceland from 1662 and later on so did Copenhagen University. In 1781 the legal deposit act was amended and since that time all printed matter was to be delivered on legal deposit (Horstboll, 1997). The first act on legal deposit stipulating legal deposit of matter printed in Iceland to Icelandic libraries was the *Act on Printing Presses no. 23/1886*, amended in 1901, 1907, 1909, 1928, and 1949. According to the amendment in 1977 published sound recordings formed a part of the legal deposit. A new legal deposit act no. 20/2002 took force the 1st of January 2003. It covers various kinds of media, amongst them publications on e-media.

Public access to published knowledge has always been considered important in Iceland. The National Library of Iceland founded in 1818 opened to the public in 1825. It was and still is located in Reykjavík, the capital. To provide comparable services in other parts of the country, the so-called Quarter Libraries, open to the public, were established in the 19th century in the west, north and east of the country. They also received legal deposits (Pétursdóttir and Júlíusdóttir, 2009). Access to official publications was brought even closer to the public by *The Act on Reading Societies and Educational Films no. 57/1937*, according to which reading societies became entitled to state support and free copies of government publications annually. Such as the parliamentary proceedings, the law book, reports published by the state and its institutions and others; provided that they counted ten or more members, submitted annual reports on their activities and income and that the local authorities in question provided at least as large a sum as provided by the state. These grants were first paid in 1939. Then there were 152 reading societies in 132 out of a total of 214 communities, serving a population of 118.888 (Statistics Iceland, 2014).

This is important in view of the fact that public access to knowledge and information is considered one of the prerequisites for functional democracy (Thorhauge, Larsen, Thun, and Albrechtsen, 1997; Larivière, 2000), and since the 19th century written knowledge, both market publications and GL, has become ever more important. Access to it is vital because no decision is better than the knowledge and information it is based on (Machlup, 1984, Vol. 3, p. 15). Apart from democratic decisions depending on access to written knowledge it is increasingly needed for every day work. The argument put forward close to a century ago that by ignoring GL the progress of science could be harmed (Auger, 1996 (1994), p. 6) is even more valid in the 21st century, when social progress depends on innovations, which in turn

^{*} Revised version of paper published in the GL15 Conference Proceedings, March 2014.

depend on access to knowledge and information; and it is less costly to use already established knowledge than having to discover and establish it anew (Georgescu-Roegen, 1999 (1971), p. 22). This applies in particular to GL in which is at the cutting edge of discoveries and new knowledge.

Previous work

The author found no studies comparable to the ones presented here. That is studies analysing the development of the whole of various kinds of knowledge output within a given country, distinguishing between market publications and GL; with a follow-up study on the effects of e-publishing on public access.

In Iceland two studies had been carried out on publications (market publications as well as GL) during the second half of the 20th century. Both focused on how the publication output developed according to the subjects published, without any regard to whether they were market publications or GL (Hjartar, 1968; Eypórsdóttir, 1997). In addition reports on market publishing in Iceland had been carried out for the Association of Icelandic Publishers (for example *Skýrsla fyrir Félag íslenskra bókaútgefenda*, 2003), not addressing GL.

Abroad studies on the importance and amount of GL in various subjects have been undertaken. There GL was estimated to be somewhere in the vicinity of 9%, towards the end of the 9th decade of the past century and to be over 20% around the middle of its 10th decade (Debachere, 1995). Other studies focused on the importance and amount of GL in research as was manifested in citations in research publications (for example Chalabi, 2012; Gentil-Beccot, 2010; MacDonald, Wells, Cordes, Hutton, Cossarini, and Soomai, 2010; Chaves, 2010).

Useful in the present studies was Wessels definition of three different types of publications with respect to bibliographic access along with access to the publications themselves. The definitions were the following: white publications, on which bibliographic information is easily accessible by the public. The publications themselves are distributed through traditional distribution channels of market publications and easily available in book stores and libraries. The second category is grey literature, which is neither bibliographically accessible in catalogues open to the public nor available through traditional distribution channels of market publishing, even when they are not confidential or secret publications. The third category is black publications. They are confidential or secret publications, which according to the definition are only meant to be accessible by the few. They are therefore inaccessible bibliographically (not found in catalogues open to the public) and the publications themselves are difficult or near to impossible to get a hold of (Wessels, 1997).

The Pisa Declaration on Policy Development for Grey Literature Resources is important for the conclusions of this study, because in it legal changes are among the recommendations for successfully providing access to GL contemporarily and for future generations (GreyNet International, 2014).

Aims of the two studies 2002 and 2012

The former study presented here was carried out in 2002, and the 2012 follow-up study carried out in 2012. Their aims and research questions differed. The aim of the former study was to analyse ways of distribution of new knowledge created in Iceland in the period 1944-2001 as it was manifested in different ratios of market publishing, GL, Icelandic patent applications and Icelandic standards. This was considered important because changes in the ratios of these kinds of knowledge outputs did affect public accessibility to knowledge and information. Access to knowledge and information distributed in market publications was easiest.

Amongst the findings of the 2002 study was an indication that e-publishing had led to decreased bibliographic access and hence decreased access to the e-publications themselves by the authorities as well as the public in Iceland. One of the reasons for this was thought to be that the *Legal Deposit Act no. 43/1977* in force at the time only stipulated deliveries of printed matter and sound recordings. Up to 1st of January when *Legal Deposit Act no. 20/2002* took force this led to a lack of deliveries of e-publications to the National and University Library of Iceland (NULI). A shortcoming that was to be amended in the 2002 *Legal Deposit Act*, which covered a greater variety of media, amongst them e-publications on internets (not those on intranets), which were to be delivered to NULI, catalogued and made accessible to the public.

The findings described above called for the second study in 2012, which was aimed at finding how access to GL had developed up to the beginning of the 21st century. In particular to find what the effects of the *Legal Deposit Act no. 20/2002* had been on legal deposit to (NULI).

Legal deposits were catalogued in Gegnir, the national union catalogue, and accessible at NULI. If not delivered properly to NULI, publications would become inaccessible bibliographically and hence hard to locate.

Research questions

The research questions in the two studies differed.

The 2002 study

The research questions in the 2002 study were the following:

1. Had there been changes in the proportion of grey versus market publications in Iceland from 1944 to 2001, and if so then which changes?
2. What was the percentage of knowledge created in Iceland and distributed as the genre of the international peer-reviewed journal article? The only source available on that issue to the researcher at the time of study was the Web of Science (WoS) database only going back to 1979.
3. Had Icelanders attempted to use Icelandic patents and standards as a way of disseminating new knowledge from 1944 to 2001, and if so then to what extent?

The 2012 study

The research questions were:

1. Whether advances in production techniques were leading to increased or diminished access to GL, both bibliographic access and access to the items themselves?
2. Whether the effects of the 2002 legal deposit act had led to increased deliveries of e-publications to NULI and hence made a higher percentage of e-GL, as well as other e-publications accessible both bibliographically and online to the public in the decade that had passed since the 2002 act took force?

Sampling strategy

The sampling strategies were different for the two studies.

The 2002 study

In the 2002 study the years of sampling were 1944, 1969, 1979, 1989, and 2001. They were chosen with regard to special events likely to have affected research and knowledge output of the nation. The assumption made was that the effects of these events on knowledge creation were realized some years after the event.

The Icelandic Research Council was founded in 1938-1940. It was the first one of its kind in the Nordic countries (Lúðvíksson, 2002). In 1944, Iceland gained complete independence from Denmark. During the forties a special department (named Atvinnudeild Háskólans), was established at the University of Iceland for research in the areas of agriculture, fishing and industry. The main function of the University of Iceland had, until then, been to prepare officials of the church and state for office, educating priests, lawyers and doctors (Hálfðanarson, Matthíasdóttir and Guðmundsson, 2011; Skúlason, 2000).

In the mid sixties the special department at the University of Iceland noted above, was discontinued and three independent research institutions established in its place by law: The Marine Research Institute, The Agricultural Research Institute and the Industrial Research Institute (*Act on Research in the Interest of Industry no. 64/1965*). At survey time, the number of research institutes, publicly as well as privately operated had increased considerably.

During the seventies considerable changes took place. Two new faculties were established at the University of Iceland, the Faculty of Social Science and the Faculty of Science (Hálfðanarson, Matthíasdóttir and Guðmundsson, 2011; Skúlason, 2000). The largest increase in the average number of publications in Iceland per year took place during that decade (*Fjöldi útgefinna bóka á íslensku*, 2006, April, 26th). The bibliography of the yearly output of publishing in Iceland that since 1944 had been published as a chapter in the *Yearbook of the National Library* (National Library of Iceland, 1945) became a separate publication in two volumes. One contained the bibliography of publications printed on paper and the other that of sound recordings. The volumes for 1979, published in 1980, were the first volumes of the national bibliography to be published using computer technology (National Library of Iceland. National Department, 1975; National Library of Iceland. National Department, 1980). That point in time marked the beginning of general use of computers in Iceland, although they had been in specialized use in the country since around 1950 (Skýrr, 2006).

During the eighties computer technology was taken into use in Icelandic libraries: catalogues in libraries in Iceland became automated and database searches became common in research and special libraries (Einarsson, 1990; Harðarson, 1993; Hauksdóttir, 1996; Júlíusdóttir, 1989, 1991, 1996).

During the last decade of the 20th century the importance of public access to information and knowledge in primary sources in the public sector was recognized as manifested in several legal acts to that effect enacted by the Icelandic Parliament (e.g. *Administration Procedures Act no. 37/1993*;

Information Act no. 50/1996; Privacy Protection Act no. 77/2000). They affected both organization of primary sources and access to them. Furthermore, Gegnir, the national union catalogue was established at the end of that decade (Hauksdóttir, 1996, 2001). The year 2001 marked the beginning of the new millennium. The last issue of *The Icelandic national bibliography* published on paper was issued (2003) for that year. Since then bibliographic information on publications issued in Iceland has been made public online in *Íslensk útgáfuskrá* (the *Icelandic National Bibliography*) (<http://utgafuskra.is>). More and more publications have been issued in the digital format and made accessible over the Internet. Web pages are moreover taking over some of the roles of newsletters. NULI established an Icelandic Web collection (www.vefsafn.is) to make them available retrospectively and for posterity (National and University Library of Iceland, n.d., 2009a, 2009b, 2011).

The 2012 study

In the 2012 study access (bibliographic and to the publications themselves) to publications of two research institutions was investigated. They were the Marine Research Institute (MRI) founded in 1965 in accordance with the Act on Research in the Interest of Industry (see above), and the Institute of Economic Studies at the University of Iceland founded in 1989. Both were chosen for their importance to the national economy and because both of them serve in an advisory function to public authorities in Iceland. Both had issued publications on paper and electronically under the 1977 and 2002 legal deposit acts and MRI also under previous legal deposit acts, and both were publishers of important GL.

Data collection

The data collection in the two studies differed.

The 2002 study

In the 2002 study the following data was collected:

1. Data on publications issued in Iceland in the sampling years. Having worked on *The Icelandic national bibliography* I knew that in spite of the time limit given in the Icelandic Legal Deposit Act on delivery of publications (*Legal Deposit Act no. 43/1977*) years could go by before some of the GL was delivered to NULI and was subsequently catalogued in *The Icelandic national bibliography*. Moreover, some of the GL, produced in-house in various institutions and organizations had not been delivered to NULI at all. I decided to use electronic data drawn from Gegnir rather than the paper version of *The Icelandic national bibliography*, in the hope that it would include grey literature such as internal reports available at organizations and institutions and catalogued in Gegnir by their staff, but never delivered to NULI and hence not found in the paper publication of *The Icelandic national bibliography*. NULI, the legal deposit library kindly, granted a copy of data on the publication output in the sampling years in electronic format from Gegnir. Using output from Gegnir rather than *The Icelandic national bibliography* on paper was expected to cover more of the GL and to give more reliable results. The data was received as an Excel file. It included: author, title, publisher, classification code, year of publication, and edition. The drawback with using the digital data from NULI was however that a tremendous amount of unanticipated work had to be put into cleaning it up in an attempt to make sure that there was only one record for each published item.
2. Data on items published on the international scholarly market. Number of items published for the international scholarly market, by authors residing in Iceland, was found by searching the Web of Science (WoS), for the sampling years as far back as the WoS went, i.e. for 1979, 1989 and 2001. The searches were carried out on the 15th of July 2002 in the *Science Citation Index Expanded (SCI-EXPANDED)* from 1970 up to July 2002, the *Social Sciences Citation Index (SSCI)* from 1970 up to July 2002, and the *Arts & Humanities Citation Index (A&HCI)* from 1975 up to July 2002 (Web of Science).
3. Information on patent applications in Iceland. Information on the number of patent applications to the Icelandic Patents Office by Icelanders in the sampling years was kindly granted for this research project by the Patents Office in Iceland.
4. Information on Icelandic standards. The information on Icelandic standards issued in the sampling years was kindly granted for this research by the Icelandic Standards Office. As it turned out standards were mostly (over 90%) adopted from abroad (Rögnvaldardóttir, 2007). Those created in Iceland proved to be too few for inclusion in this project.

The 2012 study

In the summer of 2012 data on the publishing output of the two chosen research institutions was gathered from their home pages where lists of their publications were to be found. Items on the publication list might have consisted both of GL and market publications, particularly those on the homepage of MRI.

Data analysis

Data from the two studies was analysed differently.

The 2002 study

Data from NULI, the Icelandic Patent Office, and WoS was analysed by use of the bibliometric method (Diodato, 1994; Hertz, 1987; Pritchard, 1969) in order to find out if there had been changes in the proportions of each type of knowledge output under investigation during the time of study. The data from the Icelandic Patent Office was not analyzed further according to subgenres. Calculations were made of the ratio of market publications at home and abroad (as manifested in WoS) of knowledge created in Iceland versus the ratio of such knowledge and information disseminated as GL, and as patent applications in Iceland in the sample years.

Published items were analysed as market publications when they were published by companies that had publishing for profit as their main function and as GL when published by parties that had something else than publishing for profit as their main function. My work as a cataloguer at *The Icelandic national bibliography*, made it possible for me to analyse the issuing bodies of publications in Iceland into these two types at sight. Information on knowledge created in Iceland and published abroad was only available at WoS and only for the sampling years 1979, 1989 and 2001. Calculations of the data for this 57 year period from 1944-2001 were done in Excel (Júlíusdóttir, 2006).

The 2012 study

To get an indication of the effects of the 2002 legal deposit act on how access, bibliographic as well as to the items themselves had developed, the items on the publication lists on the homepages of the chosen institutions were searched for in Gegnir.

Neither of the chosen institutions used Gegnir as their library catalogue. When publications on the lists were found in Gegnir they would therefore have been catalogued by a library in another institution most likely by NULI as legal deposit items. The searches in Gegnir were title searches with and without subtitles. Moreover keyword searches were also used when the title searches gave no results. Occasionally the year of publication in the catalogue record found in Gegnir was different from that found for the same item in the publication list on the homepage of the institution. When that happened the publication year in Gegnir was used in the analyses.

Problems and shortcomings

The 2002 study

In the 2002 study, I assumed, with regard to production techniques that information for the study was available on almost all 1944 publications, that the same was probably true for the 1969 ones, and even most of the 1979 ones. I did, however, suspect that changes in the methods of production since around 1990 (that made in-house production of publications ever easier), would have resulted in a larger percentage of publications being delivered late or not at all to NULI. Such publications would only have been found by the data collection methods used in this study, had they been catalogued in Gegnir by another institution than NULI. Their presence in the study is limited to such cases. This applies also to publications, not intended to be inaccessible or black by the publisher. Thus, one of the main shortcomings of the study as a whole was the fact that to obtain all the data and information needed was very difficult, if not impossible, due to the difficulty of finding out what had been published. This applied to GL published in Iceland in the latter sampling years and also to knowledge created in Iceland and published abroad as market publications. Therefore, the findings of a study like this one could at the best give an indication of the developments as also argued by Eypórsdóttir (1997) and Hjartar (1968).

Another shortcoming was that the 2002 study had to be limited to paper publications because before January 1st 2003 only items printed on paper and sound recordings were to be delivered on legal deposit (*Legal Deposit Act no. 43/1977*). No information was, therefore, available on e-publications in Iceland, at the time of data collection in 2002. The third major shortcoming was that it was likely that some of the 2001 publications had not yet been delivered to NULI, and that some of the delivered ones had not yet been catalogued in 2002 when the data was collected. Therefore, a smaller percentage of 2001

publications would have been available for the study than for the other sampling years. This applied in particular to GL.

The fourth major shortcoming was that when cleaning the digital data from NULI (for example deleting duplicate (or multiple) entries for the same publications), the author noticed that some publications issued in years other than the sampling years of the study were included. These were deleted but their appearance, (in the data that should only have contained records from the defined sampling years), led one to suspect that likewise some of the publications of the years under study were missing if they had not been properly input into the system.

Furthermore, it was a problem that no reliable statistics on sale of books published in Iceland existed. It was, in short neither known exactly how many titles had been issued, nor how many copies were sold of each title or in total in the country each year (Harðardóttir, personal telephone communication, June 12th 2002).

Only data in the SCI-Expanded Index, SSCI, and A&HCI at the WoS was available on the international appearance of scholarly publications by authors residing in Iceland; this limited the items worked with and did by no means cover the whole of the publications published abroad and created by authors who resided in Iceland. It was known at the time of data collection that Icelandic literature was increasingly published abroad, but statistical data on the publishing of Icelandic literature abroad proved impossible to get in 2002 from those holding their distribution rights in Iceland (Júlíusdóttir, 2006). Although not covering the output of Icelandic literature abroad by foreign publishers, it was hoped, that the study would give an indication of the developments of scientific publication activities on the international market by knowledge creators residing in Iceland.

Only patents applied for to the Icelandic Patent Office were included. Patent applications from Iceland, to patent offices abroad bypassing the Icelandic Patent Office (if any), were not included.

The 2012 study

The limited scope of the 2012 study is a shortcoming. It only covered two research institutions, whereas the 2002 study covered the whole of the publication industry in Iceland for 57 years (1944-2001), in addition to market publications abroad, although limited to those catalogued in WoS in addition to patents applications in Iceland.

Findings of the 2002 study

Table 1 provides an overview of the information sought in research questions 1-3 above. It shows that the output of written knowledge items addressed here escalated in Iceland in the period from 1944 to 2001; the increase was almost eightfold; from 404 to 3189 see Table 1, Line 7. However, the increase of each genre as a percentage of the total output differed. Thus, GL increased more than ten times while the combined market publications issued in Iceland and abroad increased almost six times, and the market publications issued in Iceland were only three and a half times more numerous in 2001 than in 1944 (see Table 1, Lines 1, 3, 4). Publications created in Iceland retrievable in WoS were under 4% of the total output of knowledge items in 1979, and over 14% in 2001 (see Table 1, Lines 2, 7). Patent applications were seven times as many in 2001 as in 1944. They were however 1.7% of the total of the written knowledge genres addressed here in 1944, but only 1.5% in 2001; while their numbers multiplied their percentage of the whole decrease in the time period under study.

The findings of the 2002 study, furthermore, indicated that the advent of e-publishing had lowered the percentages of GL catalogued in Gegnir. If that was the case, it would result in a diminished access, both public and sometimes also by the authorities to GL publications, compared to what such access was like during the era of dissemination of print on paper.

Table 1: Developments of the knowledge output in Iceland 1944-2001 of market publishing, grey literature, the genres of international peer-reviewed journal articles and patent applications to the Icelandic Patent Office

Year of publication	1944	1969	1979	1989	2001	Line
No issued by commercial publishers	204	357	558	884	727	1
No issued by commercial publishers abroad, WoS	NA	NA	64	122	452	2
Total issued by commercial publishers	204	357	622	1006	1179	3
No issued by grey literature publishers	193	530	1021	862	1961	4
Total issued	397	887	1643	1868	3140	5
No of patent applications	7	26	9	18	49	6
Total output in Iceland inc. WoS	404	913	1652	1886	3189	7
						9
% issued by commercial publishers	50.5	39.1	33.8	46.9	22.8	10
% issued by commercial publishers abroad, WoS	NA	NA	3.9	6.5	14.2	11
% issued by commercial publishers in total	50.5	39.1	33.6	53.3	37.0	12
% issued by grey literature publishers	47.8	58.1	61.8	45.7	61.5	13
% of patent applications	1.7	2.8	0.5	1.0	1.5	14
% of total in DDC 800 (literature) including WoS	37.0	31.0	25.0	25.0	19.0	15

Already in 1944 GL accounted for 47.8% of total publishing in the country and was up to 61.8% of the total publishing in 1979. Since then GL-items as a percent of the whole of written knowledge catalogued in Gegnir in 1989 decreased and were only up to 61.5% in 2001 (see Table 1, Line 13) (Júlíusdóttir, 2006, 2013). One of the reasons was that the publishing output of market publishers in 1989 was unusually high. This finding was confirmed both by the editor of the *Icelandic national bibliography* and the manger of the Icelandic Publishers Association.

Moreover the percentage of literature (classified in 800 in *Dewey decimal classification system* (DDC)) decreased steadily. It went from 37% in 1944 of the total knowledge output to 19% in 2001 (see Table 1, Line 15). At the turn of the century the total numbers of publications had increased much more than the number of the inhabitants (see Figure 1). This is considered to be an indication of increased dependence of the working culture of Icelandic society on written knowledge; and that research was becoming a way of making a living for many people in Iceland. Not to witness an increase since 1979 in the percentages of the output of GL at the turn of the 21st century, when research activity in the country had increased greatly could not be right. It is argued here that this was an indication of in-house production of GL (for example research reports and other kinds of GL) and as the century drew to a close increased use of e-GL which was not covered by the legal deposit act and therefore probably not catalogued in Gegnir (see

Figure 1 and Table 1). Standards created in Iceland were too few for inclusion in the study (see above).

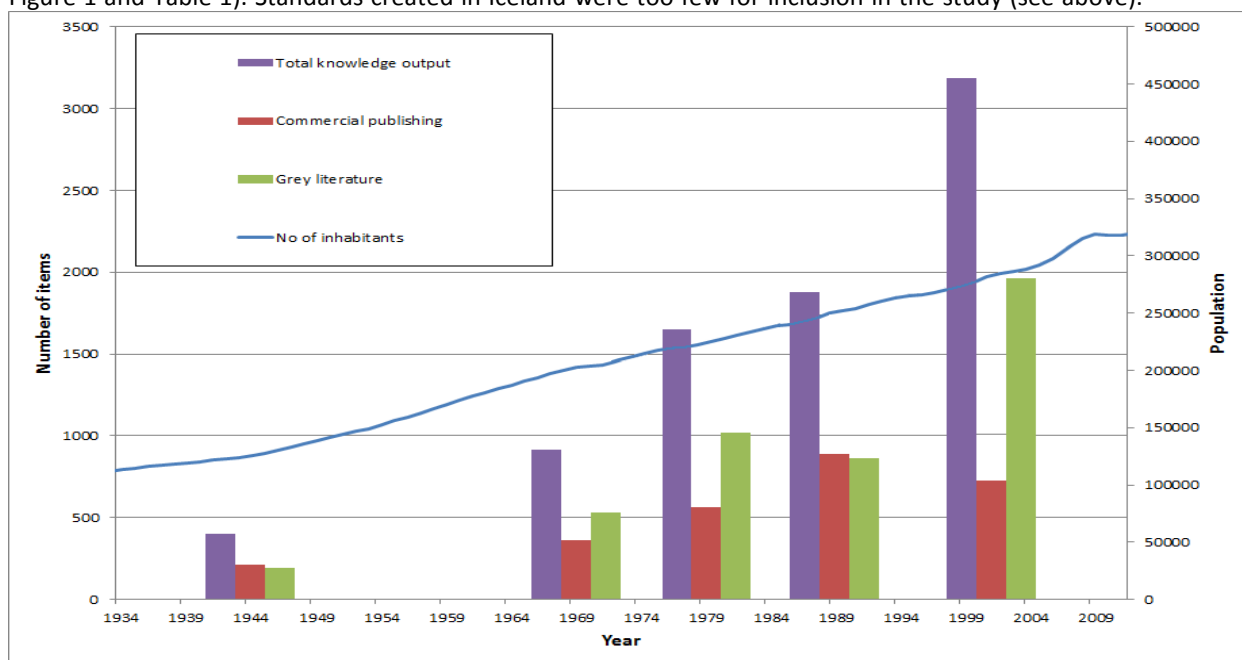


Figure 1: Development of total knowledge output in Iceland, market publications (issued in Iceland and abroad found in WoS), GL and patent applications in Iceland are depicted in the columns. The line shows the increase in the population for comparison.

Findings of the 2012 study on the effects of the Legal Deposit Act no. 20/2002 on access to GL

The findings of the 2012 study indicated that the advent of e-publishing had lowered the percentages of their publications catalogued in Gegnir; probably because they had not been delivered to NULI. Resulting in a lack of public access and sometimes also of access by the authorities to these publications, even though they might be open on the Internet on the homepages of their issuing bodies, proper bibliographic access was needed to locate them (see Figures 2 and 3). Since the data collection was carried out in the summer of 2012 a few publications issued by the two institutions under study had been catalogued in Gegnir (as found in searches on November 14th 2013).

Figures 2 and 3 give an indication of legal deposit deliveries to NULI and show national bibliographic visibility of the items that had been catalogued in Gegnir at the time of study. It also shows Internet access to the full text available on the home pages of the institutions of study. At MRI (<http://www.hafro.is>) all items on the publication list on its homepage had been scanned and were available as pdf documents to the public (see Figure 2).

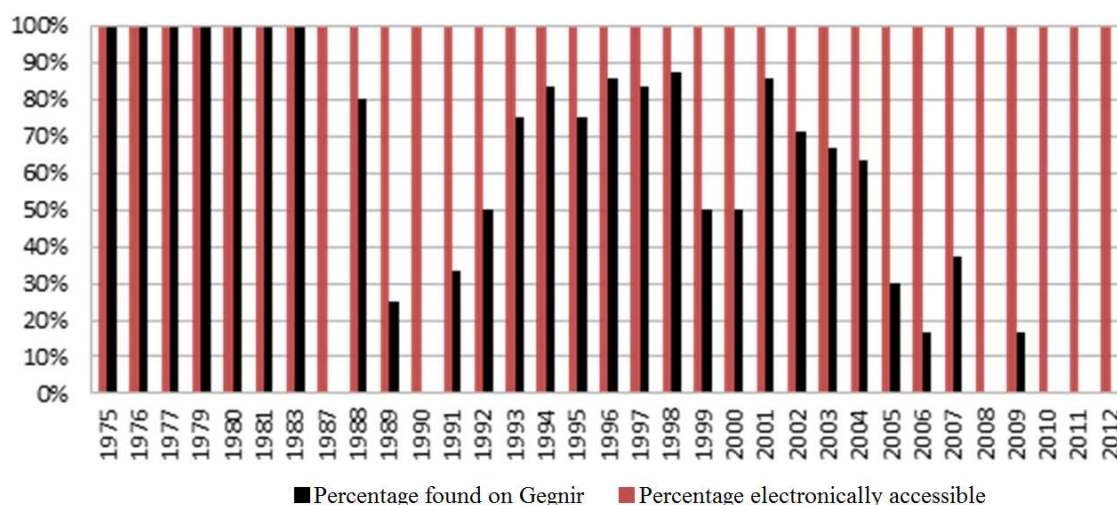


Figure 2: Percentages of publications of the Marine Research Institute catalogued in Gegnir in the summer of 2012 and thereby available bibliographically and physically for use by the public.

A similar kind of access was available to some of the publications on the list of the Institute of Economic Studies (<http://hhi.hi.is/>) although the e-items there were mostly items published in this century and therefore produced electronically (see Figure 3).

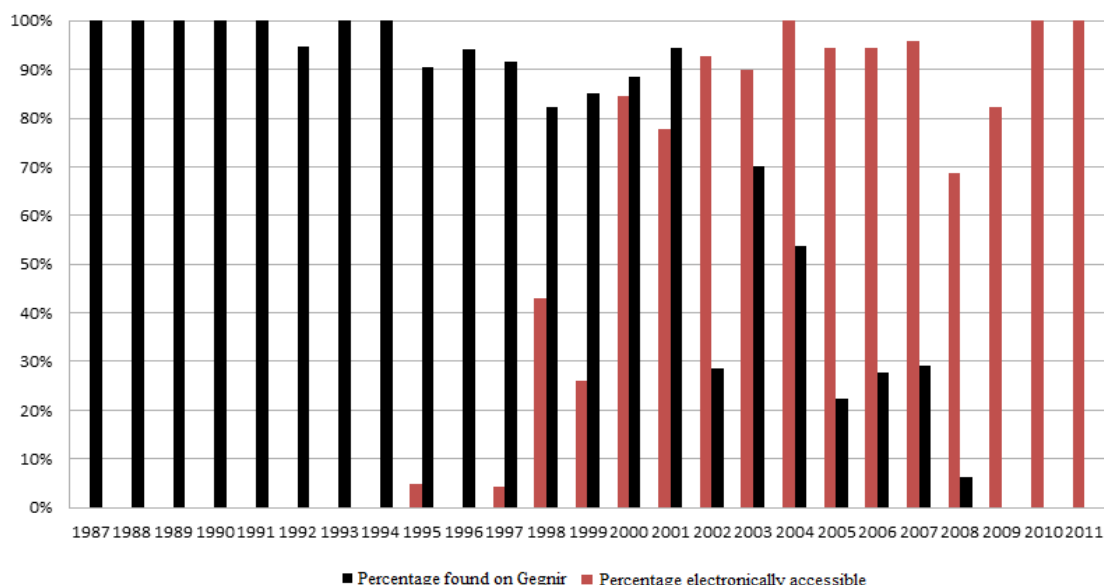


Figure 3: Percentages of publications of the Institute of Economic Studies at the University of Iceland catalogued in Gagnir in the summer of 2012 and thereby available bibliographically and physically for use by the public.

Contrary to expectations, bibliographic access in Gagnir to GL issued by the two institutions investigated did not improve after the 1st of January 2003. This meant that finding items known to have been published, when the subjects terms did not appear as words in their titles could be almost impossible. Google searches proved imprecise and even useless in some such cases. So that even when e-GL publications might be open on the Internet on the institutions' homepages, they could not be found due to lack of bibliographic cataloguing (see figures 2 and 3).

When it was known which institution had published the e-item in question, a user would have to open and read a great many items on a long list of publications to find out which item was the right one. This was very time consuming to say the least.

According to staff at NULI there was not a backlog of publications waiting to be catalogued; such a situation would have made matters right in the end, when all had been catalogued. The only explanation remaining was that the publications of these important institutions had simply not been delivered to NULI. Among the reasons for this development was that the staff of the two institutions claimed not to have been aware of their legal duty to deliver their publications to NULI. It is worth noting in this context that there had been staff changes in both institutions.

Discussion

The finding of the 2002 study that close to half of all publications catalogued in Gagnir were GL in Iceland in 1944, and the high GL percentage of the whole of knowledge output indicated that legal deposit deliveries were in relatively good order. The Icelandic GL situation certainly was quite different from that in some other countries, judging by the estimates abroad that GL made up 9% of the publishing by the end of the 9th decade and that this had risen to over 20% by the middle of the 10th decade of the past century (Debachere, 1995). It would seem that these percentages were bound to be an underestimation when compared to the percentage of GL in Iceland, which was a rather backwards country in 1944 with few research institutions and only one university, which mostly had the role of educating public officials. When that situation changed the percentage of GL went up above 60% in 1979.

The increases in the numbers of higher educational institutions and increased research activity in the country, is manifested in the increase in GL (although less than expected after 1979), which indicated that GL was increasingly produced in-house and electronically as the 20th century drew to a close. There were also increases in the percentages of market publications created in Iceland and published abroad as retrieved in WoS, and in patent applications.

The aim to provide the public with access to publications has been clear regarding access to both GL and market publications since the opening of the National Library to the public in 1825. This aim has, moreover been manifested in various legal acts on written and published knowledge and information. Up to the time when technological developments came to facilitate production of publications directly by the issuing bodies, all publications in Iceland, apart from the black were white according to Wessels definitions (1997), in the sense that they were catalogued in *The Icelandic national bibliography* and available for consultation by the public at The National Library in Reykjavík and the three Quarter Libraries, and quite possibly also at the issuing bodies themselves. Users would have had to go these places to get hold of them. The advent of e-publishing was expected to make access to GL physically easier through dissemination on home pages of the issuing bodies (Farace and Schöpfel, 2010; Wessels, 1997); in particular the access to contemporary GL (Júlíusdóttir, 2006). It could hardly be expected that the same GL would be accessible on the home pages forever. To secure public GL-access of posterity NULI established a web collection (vefsafn.is) of chosen web-pages, as noted above. These findings show, however, that lack of proper bibliographic access, is a major hindrance to contemporary GL-access in some cases, to say nothing of access by posterity in such cases.

Contrary to expectations, access to GL issued in Iceland did not improve subsequent to the *Legal Deposit Act no. 20/2002* taking force January 1st 2003. One of the reasons why staff of the issuing bodies of GL was unaware of their duty to deliver their publications to NULI, may be that NULI had not informed sufficiently on their duty to deliver legal deposit copies. Although the 2002 act covered a greater number of kinds of media than the 1977 act, its scope is narrower than that of previous such acts in Iceland, in the sense that in the area of e-publishing it only stipulates legal deposit of internet publications and not intranet publications. This results in a situation where the same publication is to be delivered on legal deposit if published on paper or disseminated on the internet, but not if disseminated on an intranet. In Iceland publications for example on research work in progress are typically disseminated as GL during developmental stages many a time only on intranets and the final results are published on the international scholarly market, and sometimes not in Iceland at all. The effects of the *Legal Deposit Act no. 20/2002* are that entire research projects might be carried out in Iceland without delivery of and hence access to any publication on them being available in Iceland to authorities and the public presently or in the future.

Further research

Further research is needed to establish the effects of the stipulations in the 2002 act regarding access to e-GL. Moreover, the visibility and access to knowledge created in Iceland and published abroad has to be investigated.

Conclusions

The development described above calls for a revision of the *Legal Deposit Act no. 20/2002*, a recommendation in accordance with The Pisa Declaration on Policy Development for Grey Literature Resources (GreyNet International, 2014). The revision should be based on the traditional spirit of access of the public and authorities to knowledge created in Iceland and to publications issued in Iceland, GL as well as market publications. As this was written some e-repositories were operated in the country. Rafhladan.is is an e-repository operated by NULI covering a wide subject area and GL (National and University Library of Iceland, n.d.). Examples of other repositories are the institutional repositories skemman.is for The Agricultural University of Iceland, The Iceland Academy of the Arts, The National and University Library, The University of Akureyri, The University of Bifröst, The University of Iceland and The University of Reykjavik. Skemman.is houses students' digital theses and dissertations in addition to articles and other research material from the universities' academic staff. Hirsli.is is a repository for The National University Hospital. The coverage of these repositories is far from including all of the knowledge created in Iceland and published there and abroad. Moreover, such repositories are far from providing the bibliographic access needed and from covering all the GL that should be available for access in Iceland. All Icelandic publications have to be catalogued bibliographically and be linked to the national union catalogue as a means of locating them to be most useful. Rules have to be set on which items to make thus accessible and to preserve for the future as it is highly unlikely that all can be maintained for future use, as was the case when fewer items were published.

References

- Auger, C. P.** (1996 (1994)). *Information sources in grey literature* (3. edition). London: Bowker.
- Chalabi, Lydia.** (2012). Characteristic and use of grey literature in scientific journals articles of Algerian Researchers: Case study of University of Science and Technology Houari Boumediene (Physics, Chemistry and Computer Sciences). In Farace, Dominic and Frantzen, J. *GL14 Conference Proceedings*. Amsterdam: TextRelease, 2013.
- Chaves, Todd A.** (2010). Grey literature in Karst research: The evolution of the Karst information portal (KIP). In Farace and Schöpfel (editors), *Grey literature in library and information studies* (pp. 181-197). Berlin: Walter de Gruyter GmbH & Co. KG.
- Debachere, M-C.** (1995). Problems in obtaining grey literature. *IFLA journal*, 21 (2), 94-98.
- Diodato, Virgil P.** (1994). *Dictionary of bibliometrics*. New York: Haworth Press.
- Einarsson, Eiríkur Þ.** (1990). Nýting erlendra gagnabanka. [Use of foreign bibliographic databases]. In *Upplýsingar eru auðlind* (pp. 259-269). Reykjavík: STUPP.
- Eypórsdóttir, Hildur G.** (1997). Íslensk bókaútgáfa í 30 ár, 1967-1996. [Icelandic publishing during 30 years, 1967-1996]. *Ritmennt*, 2, 76-80.
- Farace, Dominic J. and Schöpfel, Joachim (Eds.).** (2010). *Grey literature in library and information studies*. Berlin: De Gruyter Saur.
- Fjöldi** útgefina bóka á íslensku. (2006, April 26th). [Number of publications in Icelandic]. *Landskerfi bókasafna. Fréttasafn*. Retrieved May 12th 2006 at: <http://www.landskerfi.is/sja.php?id=206>.
- Gentil-Beccot.** (2010). The driving and evolving role of grey literature in high-energy physics. In Farace and Schöpfel (editors), *Grey literature in library and information studies* (pp. 155-166). Berlin: Walter de Gruyter GmbH & Co. KG.
- Georgescu-Roegen, Nicholas.** (1999 (1971)). *The entropy law and the economic process*. Cambridge, Mass.: Harvard University Press.
- Grey Net International.** (2006). *Grey literature network service*. Retrieved August 3rd 2006 at: <http://www.greynet.org/>.
- Grey Net International.** (2014). *The Pisa Declaration on Policy Development for Grey Literature Resources*. Retrieved July 24th 2014 at: <http://greyguide.isti.cnr.it/include/pisadeclarationmay2014.pdf>.
- Hálfðanarson, Guðmundur, Matthíasdóttir, Sigríður and Guðmundsson, Magnús.** (2011). *Aldarsaga Háskóla Íslands 1911-2011*. [Centennial history of the University of Iceland]. Reykjavík: Háskólaútgáfan. Karlsson, Gunnar (Ed.).
- Hauksdóttir, Sigrún** (2001). *Consortium of Icelandic libraries*. Retrieved February 14th 2006 at: <http://www.bokis.is/iod2001/papers.html>.
- Hauksdóttir, Sigrún.** (1996). Gegnir. [Gegnir: The automated library system of the National and University Library]. *Bókasafnið*, 20, 64-67.
- Harðarson, Hrafn.** (1993). Tölvuvæðing í bókasöfnum á Íslandi. [Automation in Icelandic libraries]. *Bókasafnið*, 17, 17-22.
- Hertz, Dorothy.** (1987). Bibliometrics, history of the development of ideas. In *Encyclopedia of library and information science* (Vol. 42, supplement 7, pp. 144-219). New York: Marcel Dekker.
- Hjartar, Ólafur F.** (1968). Íslensk bókaútgáfa 1887-1966. [Publishing in Iceland 1887-1966]. In *Landsbókasafn Íslands. Árbók 1967*. [National Library of Iceland. Yearbook 1967].
- Horstboll, Henrik.** (1997). *The Royal Library over 350 years* (Vibeke Cranfield, Trans.). Copenhagen: The Royal Library.
- Júlíusdóttir, Stefanía.** (2013). *A role to play: Continuity and change in career opportunities and working conditions in libraries, records management, and archives*. Reykjavík: University of Iceland. PhD dissertation.
- Júlíusdóttir, Stefanía.** (1989). Plans for co-operative automation in Icelandic libraries. *Scandinavian Public Library Quarterly*, 22(1), 21-26.
- Júlíusdóttir, Stefanía.** (1991). The state of automation in Icelandic libraries: Past, present and future: An overview. In *OPAC's and the User: Proceedings of the third Anglo-Nordic Seminar* (pp. 59-62). Esbo: Nordinfo.
- Júlíusdóttir, Stefanía.** (1996). Tölvuvæðing íslenskra bókasafna: Könnun á stefnumörkun og stöðu mála. [Automation in Icelandic libraries: Policies and practices]. *Bókasafnið*, 20, 54-63.
- Júlíusdóttir, Stefanía.** (2006). Changes in publishing and LIS work in Iceland. In Hauksdóttir, Sigrún (Ed.), *Rannsóknir í félagsvísindum VII: Félagsvísindadeild* (pp. 65-78). Reykjavík: Félagsvísindastofnun Háskóla Íslands.
- Larivière, Jules.** (2000). *Guidelines for legal deposit legislation* (a revised, enlarged and updated edition of the 1981 publication by Jean Lunn IFLA Committee on Cataloguing). Paris: IFLA. Retrieved January 4th 2012 at <http://archive.ifla.org/VII/s1/gnl/legaldep1.htm>.

Lúðvíksson, Vilhjálmur. (2002). [Presentation given at the Annual Meeting of the Icelandic Research Council, by Vilhjálmur Lúðvíksson the Managing Director of the Icelandic Research Council]. Paper presented and distributed on paper at the Annual Meeting of the Icelandic Research Council 2002, Reykjavík.

MacDonald, Bertrum H., Wells, Peter G., Cordes, Ruth E., Hutton, Gregory R.G., Cossarini,

Danielle M., and Soomai, Soonai S. (2010). The Use and Influence of Information

Produced as Grey Literature by International, Intergovernmental Marine Organizations:

Overview of Current Research. In Farace and Schöpfel (editors), *Grey literature in library and information studies* (pp. 167-180). Berlin: Walter de Gruyter GmbH & Co. KG.

Machlup, Fritz. (c1984). *The economics of information and human capital* (Vol. 3). Princeton: Princeton University Press.

National and University Library of Iceland. (2009a, July). *Fréttir 29.7.2009: Vefsafn opnað fyrir prófanir*. [News 29.7.2009: Web collection opens for trial]. Retrieved April 27th 2011 at: <http://landsbokasafn.is/index.php/news/128/15/Vefsafn-opnad-fyrir-profanir>.

National and University Library of Iceland. (2009b, September). *Fréttir 29.9.2009: Vefsafn opnað*. [News 29.7.2009: Web collection opens]. Retrieved April 27th 2011 at: <http://landsbokasafn.is/index.php/news/135/15/Vefsafn-opnad>.

National and University Library of Iceland. (n.d.). *Rafhlaðan*. [Collection of electronic publications issued in Iceland]. Retrieved May 25th 2013 at: <http://rafhladan.is/>.

National and University Library of Iceland. (2011). *Vefsafn*. [Web collection]. Retrieved April 27th 2011 at: <http://vefsafn.is/>.

National Library of Iceland. National Department. (1975). *Icelandic national bibliography 1974*. Reykjavík: National Library of Iceland.

National Library of Iceland. National Department. (1980). *Icelandic national bibliography 1979*. Reykjavík: National Library of Iceland.

Pétursdóttir, Kristín and Júlíusdóttir, Stefanía. (2009). The history of public libraries in Iceland. In Dyrbye, Makinen, Reimo and Torstensson (editors), *Library spirit in the Nordic and Baltic countries: Historical perspectives* (pp. 13-26). Tampere: HIBOLIRE.

Pritchard, Alan. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25(4), 348-349.

Rögvaldardóttir, Guðrún. (2007). Staðlaráð tvítugt, eða þannig: Brot úr sögu Staðlaráðs. [The Standards Organisation becomes twenty]. *Staðlamál*, 11(3), 2.

Sigurðsson, Einar. (1971). *Saga íslenskrar prentunar og bókaútgáfu: Ágrip*. [History of printing and publishing in Iceland] (2nd edition). Reykjavík: [s.n.].

Skúlason, Páll. (2000). *Vísindi og fræði á aldamótaári*. [Science and research at the millennium. Presentation delivered on January 26th 2000]. Retrieved January 26th 2000 at: <http://www.visindiogfraedi.hi.is/pallsk.html>.

Skýrr. (2006). *Skýrr*. Retrieved July 20th 2006 at: <http://www.skyrr.is>.

Skýrsla fyrir Félag íslenskra bókaútgefenda: Áfangaskýrsla. (2003). [A Report for the Association of Icelandic Publishers: An interim report]. Reykjavík: Institute of Economic Studies at the University of Iceland.

Statistics Iceland. (2014, 03.04). *Population: Key figures 1703-2014*. Retrieved March 10th 2014 at <http://www.statice.is/Statistics/Population/Overview>.

Thorhauge, Jens, Larsen, Gitte, Thun, Hans-Peter and Albrechtsen, Hanne. (1997). *Public libraries and the information society. Luxembourg*: European Commission.

Wessels, R. H. A. (1997). Het belang en toegankelijkheid van grijze literatuur [The importance and accessibility of grey literature]. *Informatie professional*, 1(3), 28-30, 32.

Scholarly publishing behaviour in Slovakia ... Are we ready for repositories?*

Jaroslav Šušol, Miriam Ondrišová, Jana Ilavská, and Marcela Kopecká
(Slovak Republic)

Abstract

Institutional repositories are regarded as one of major conceptual and technological revolutions in the field of scholarly publishing, enabling quick and widespread dissemination of research results. In order for the repositories to become this useful and powerful tool, it is necessary that they get full political and financial support from the academic institutions' managements as well as a wide acceptance from the academics, university teachers and researchers.

At the Comenius University in Bratislava, the oldest and the largest university in Slovakia, we are preparing a project of new institutional repository. In order to find out about the prospects of such an endeavor, we plan to make an enquiry on its feasibility from both the above-mentioned points of view – 1. state of the art of repositories implementation in Slovakia and 2. the acceptance of electronic publishing channels within the complex of manifestations of publishing behavior.

One of the principal approaches towards research in publishing behavior is the application of quantitative methods of scientometrics or bibliometrics that are usually put into a broader context. Scientometrics is not a new discipline, but very current at the moment in Slovakia, vividly discussed in the academic circles due to the fact that some of its methods and indicators are being recently used as an important part of university management and of allocation of financial resources from the level of ministry of education.

We will apply both sociological and scientometrics methods on publishing data from Comenius University in our research. It should lead to defining the patterns of publishing behavior of authors in the academic environment, with respect to the particularities in humanities and social sciences, natural sciences and technical disciplines.

Keywords: institutional repository, publishing behavior, electronic publishing, online publishing, bibliometrics, scientometrics

1. Introduction

Research data and research papers are frequently cited as some of primary examples of grey literature. Their accessibility is often limited to a low number of copies, especially, of course, in case of printed volumes. Institutional repositories are regarded as one of major conceptual and technological revolutions in the field of scholarly publishing, enabling quick and widespread dissemination of research results. In order for the repositories to become this revolutionary tool, it is necessary that they get full political and financial support from the academic institutions' managements as well as a wide acceptance from the academics, university teachers and researchers.

At the Comenius University (CU) in Bratislava we have started a project of building a scientific park in the domain of molecular medicine, environmental medicine and biotechnologies in 2013. Several supporting project activities are devoted to social and humane aspects of biotechnologies (legislation, ethics, economics). One of these socio-cultural streams of research is focused on analysis of users and usage contexts of institutional repositories and creation of the CU institutional repository model.

In the framework of this research goal we plan to make an enquiry on its feasibility from both the above-mentioned points of view, and to look at:

1. state of the art of repositories implementation in Slovakia,
2. the acceptance of electronic publishing channels within the complex of manifestations of publishing behavior, including the researchers' willingness and readiness to share data and resources.

In this paper we present the results of some of our pilot research efforts into the domain of publishing behavior and repositories in Slovakia. Preliminary characteristics of publishing culture/reality in different scientific disciplines, or rather broader discipline areas, were compiled on the basis of bibliographic data contained within CU Library database of CU employees' publications. Application of quantitative methods of scientometrics or bibliometrics can provide us with an exact picture of distribution of individual bibliographic parameters within the database and define some of the features of publishing behavior at Comenius University.

* First published in the GL15 Conference Proceedings, March 2014.

2. Institutional repositories in Slovakia

Institutional repository is a complex, multi-structured device whose primary role lies in collecting, archiving and disseminating digital publications and data coming from a particular, usually institutionally defined environment. This phenomenon is closely connected with such concepts and technologies as scholarly communication and publishing, digital library and open access.

Traditional approach to classification of repositories takes into account various aspects of their content and its geographical limits, recognizing e.g. subject-based repositories, research repositories, national repositories or institutional repositories. Predominantly, building of institutional repositories is connected with academic institutions where repositories are designed to hold the full-texts of the complete intellectual output of institutions' individual members. In this respect, institutional repository fulfills one more important role – the role of presentation, marketing the institution in the digital space. Institutional repositories and open access (OA) can have various practical manifestations, depending mostly on the nature of material that is being archived and disseminated which, consequently, has an impact on its economic and, possibly, legislative functioning. Two principal streams of open access publishing have been defined as the Gold model and the Green model (Craig 2007, Houghton 2013). The Gold model is based on a traditional journal as a vehicle for publication, with possible changes in the financial model, which is dominantly saturated by authors or sponsors. Within the Green model, the authors usually post their manuscript into an electronic archive (institutional or subject-based), as a pre-print (in the form that was submitted to a journal) or as a post-print (a final copy of the peer-reviewed article).

From the very beginning of the development in the area of online publishing and repositories it was clear that one of the main advantages of these new technologies was the increased visibility, or rather, better findability of the material and publications that were published digitally on the net. Some of the early research results in this area (Lawrence 2001, Harnad 2004) even showed that there is a correlation between free online availability of publications and higher citation counts.

More recent studies do not see this relation so clearly. Some authors have identified three possible reasons that could account for the observed citation differences between OA and non-OA articles. They were termed an *Open Access postulate* (authors are more likely to read, and thus cite, articles available freely), a *Selection Bias postulate* (the most prominent, and thus most citable, authors are more likely to make their works freely available), and an *Early View postulate* (the period between the early posting of an article, either pre-print or post-print, and the appearance of the cognate published journal article allows for earlier accrual of citations) (Craig 2007).

We looked through the state of building of the institutional repositories in Slovakia via questionnaire survey within research project VEGA 1/0429/10 *Information ecology of the academic information environment* (Steinerová 2012). Questionnaire had been addressed to all Slovak universities, colleges and the Slovak Academy of Sciences (total number of addressees - 43). We received 27 completed questionnaires (response rate 69.2%). All significant academic institutions had been among respondents. Our findings were not very positive, only 14 institutions reported to have some form of institutional repository - 2 institutions have been building standalone institutional repository and 13 institutions are using local storage with web access to cover some of repository functions (one of the institutions has both, standalone repository as well as local storage). Twelve institutions have not any kind of repository yet, but they are planning to build it in relatively short time. One institution has neither repository nor plans to build it. It is interesting, that in majority of cases (18) it was the library (often in cooperation with IT department) which initiated or is initiating the creation of repository within institution, while administration and management did not show a lot of interest in this matter. The survey detected the most common problems and barriers that institutions used to face, among them especially copyright and other related legislation, lack of cooperation among relevant departments within institution (management, library, IT department, archive, publishing house, etc.), reluctance to share research data out of official and reputable publication channels, individual traits (computer skills, age).

In Slovakia, the state of the art of repositories was predominantly influenced by the Act on Higher Education No. 131/2002, and its subsequent amendments. Since 2010, due to the efforts connected with building a system of plagiarism/originality control, all the universities in Slovakia are obliged to collect all qualification works and final theses (bachelor, master, dissertation, etc.) in electronic format and send them into a state-wide Central Register of Theses (<http://www.crzp.sk>). Similarly, each university has to build a database focused on bibliographical registration of publications of its employees. This database, however, is not supposed to contain digital copies of the publications (full texts). Again, all the data from individual databases are merged into the Central Register of Publications (<http://www.crepc.sk>).

These two types of registers that are being built at all universities in Slovakia can be seen as a good starting point in all efforts and plans to develop institutional repositories. Although the register of publications at each university usually contains only bibliographic descriptions of books and articles published by university researchers, each academic library is obliged to collect hard copies of these documents (full text in case of articles, principal pages in case of books, proceedings and collections of papers). So, the organizational system of getting hard copies of documents is in place, the task now should be to turn it into digital environment.

3. Publishing behavior

Publishing behavior represents an important part of *information behavior* of information users. While information behavior in the broadest sense is usually defined as a complex of various ways of interaction between a human as an actor of information processes and the information, the publishing behavior is a specific component of this phenomenon, that occurs on the production side of information and communication spectrum. It is embodied in various ways and approaches that the authors apply when it comes to publishing, and especially to selecting and using diverse communication and publication channels.

It is possible to explore the publishing behavior using various approaches, quantitative and qualitative methods, analyzing the subjective interpretation of reality among authors themselves (sociological or ethnological methods based on questionnaires, surveys or interviews) or measuring hard data in the form of publication outputs, production. Although there are researchers in this area who like to stress that there are some methodological drawbacks to these enquiries. For example, Michels and Schmoch (2013) stipulate that "it has been confirmed in the social sciences that subjects under observation change their behavior, making the genuine observation of their behavior almost impossible. This phenomenon, called Hawthorne effect, is more pronounced in cases where the results of the observation have consequences for the observed subjects".

One of the most important impulses for the development of publishing behavior research arrived in 1960s, with the application of computing technology into the processes of secondary information storage and retrieval. Research projects concentrated mostly on specific conditions and problems of particular scientific discipline (Cox 1993, Luukkonen 1992, Swan 2008, Kyvik 2003, Rey 1998, Rockwell 2000, Björk 2000 etc.).

The situation in economics since 1970s was analyzed by Cox (1993). He presented various aspects that can be studied in connection with the production and distribution of economics literature, like contributions of individual institutions to the economics journals and their impact on the development of their reputation, searching for „doctoral roots“ of individual authors, relative quality of journals and their influence on economics literature. As Cox pointed out, „the knowledge of the publication behavior of individuals can serve many useful purposes. For instance, if it is assumed that this behavior of individuals is stable over time, it will help assess the likelihood of multiple publications by individuals in the future. In addition, such knowledge provides valuable information about the implicit strategies of economic researchers in selecting their research outlets, the behavioral aspects of joint work (i.e., co-authored paper), and the relationship between the publication count and an author's stage in the career cycle.“ (Cox 1993).

At the beginning of 1990s, Luukkonen (1992) sought to verify the assumption that in their publishing behavior the scientists tend to select prestigious publishing channels, seeking an optimal level in the hierarchy of publications. The interviews with researchers in three areas (zoology, biomedicine, automation and control technology) suggested that when it comes to selecting the publishing channel, the Finnish scientists value equally the criteria of reward and communication functions of publishing. The real publishing behavior of biomedical scientists, however, shows their leaning towards higher appreciation of publication prestige (Luukkonen 1992). More recent research from Great Britain suggests that the majority of scientists prefer the journal titles with relevant reader community. This choice is not always in line with the highest impact factor journals in a given discipline. Publishing in journals with high impact factor usually brings advantages in the form of higher formal criteria research assessment, however, the reach to the right audience brings reward in terms of recognition by peers (Swan 2008).

There were also attempts to map the development of publishing behavior on a wider interdisciplinary scale. Kyvik's research focused on assessing the evolution of publishing behavior within 2 decades and analyzed changes in publication patterns at the universities in Norway. Based on 3 questionnaire surveys from 1982, 1992 a 2001 he concluded that co-authorship became a more common way of publishing, the extent of publishing oriented towards an international audience has increased, the scientific article in an international journal consolidated its position as a dominant type of publication and the number of

publications per academic researcher increased. One of the most intriguing conclusions states that the publication patterns in social sciences get closer to those in natural sciences (Kyvik 2003). A similar research, although restricted to the field of agronomy, was carried out in Spain. The analysis of articles published in national and international journals between 1980 and 1995 revealed 2 leading tendencies – migration of works towards SCI journals and increased use of books and monographs as channels of publication of research works (Rey 1998).

In 2000, trustworthiness of electronic publishing was put into focus in the research that took place at the universities in Canada. The results showed that only 16% of respondents published their articles in a peer-reviewed electronic resource, while as many as 61% used other, traditional communication channels (Rockwell 2000). The Swedish survey from the same year concentrated on the situation in building engineering and concluded that both the students and professors consider reviewing to be an important part of electronic publishing processes. When comparing traditional and electronic regimes of publishing, 75% of respondents said that the speed of publishing on the web was more important for them than sticking to formal publishing procedures (Björk 2000).

The literature, however, is not always absolutely positive about electronic publishing, it also reflects the drawbacks of using new technologies for the purposes of scholarly publishing. As the Nature points out – „Scientific productivity, as measured by scholarly publication rates, is at an all-time high. However, high-profile cases of scientific misconduct remind us that not all those publications are to be trusted – but how many and which papers? Given the pressure to publish, it is important to be aware of the ways in which community standards can be subverted. Our concern here is with the three major sins of modern publishing: duplication, co-submission and plagiarism.“ (Errami 2008).

First major survey of this type in Slovakia was done in 2002 by means of a questionnaire among the users of academic libraries. It focused, among other things, on the authors' relationship towards traditional and electronic publishing technologies. The results supported the assumption that the acceptance of electronic communication channels for publishing the output of scientific research depends on particular discipline, with strong unbalance between social sciences and humanities on one side and natural and technical disciplines on the other (Šušol 2004).

A combination of qualitative and quantitative methods was applied in the survey of publishing preferences of PhD students at the Faculty of Arts, Comenius University in Bratislava. One of the key issues was the authors' attitude towards traditional and electronic publishing. The conceptual analysis of the background of individual author's attitudes or choices, revealed 3 most frequent categories – *time*, *availability* and *reputation*. First 2 categories (time and availability) are being mostly connected with electronic publishing. The results also indicate that the refusal of electronic publishing channels is caused rather by their non-acceptance in the authors' closest peer environment, not that much by the fact that the authors themselves would not realize the overall circumstances, advantages and disadvantages of electronic publishing (Steinerová et al. 2006).

The above-mentioned examples and results demonstrate that the publishing behavior is a multi-dimensional research domain, as the behavior of information process actors is influenced by various factors. The attitudes, opinions, preferences, that together define publishing behavior of authors in academic area, are not only subject to evolution in time, but have also certain local or national particularities.

One of the principal approaches towards research in publishing behavior, of course, is the application of quantitative methods of scientometrics or bibliometrics that are usually put into a broader context. We will eventually apply both sociological and bibliometrics methods on publishing data from Comenius University in our research. It should lead to defining the patterns of publishing behavior of authors in the academic environment, with respect to the particularities in humanities and social sciences, natural sciences and technical disciplines. There are several variables that can be calculated and compared in this way, e.g.:

- publishing in foreign language / Current Contents Connect (CCC) journals
- usage of foreign (or foreign language) resources / electronic resources
- types of publications
- co-authorship
- methods and quantity of citing, referring
- attitude towards publishing in the network environment / open access regime.

At the moment, though, we will only present results of some of our preliminary analyses, aimed at finding connections between electronic publications, or rather electronic/online versions of publications on one side, and their availability and citation rate, on the other.

4. Comenius University database of publication outputs analysis

Bibliographical registration of publication outputs of academic staff in Slovakia is regulated by Library Act No. 183/2000 Coll. and Act No. 455/2012 on Higher Education Institutions and Directive of Ministry of Education No. 456/2012 Coll. Registration process distinguishes between different types of publications (monograph, chapter, journal article, review, conference paper, proceedings, etc.) and uses different criteria (originality, territoriality, communication channel, target audience, etc.) to categorize publication outputs within overall 83 categories. Some of the categories are considered to be of higher quality or value (e.g. A – scientific works), as they publish the results of original scholarly research, usually validated by the quality of publication channel (CCC journal, reviewed collection of papers etc.). Bibliographic processing of publications includes also registration of citations to the publications (excluding authors' self-citations), there are 8 categories of citations (e.g. foreign/domestic registered in citation indexes in Web of Science or Scopus, not registered in citation indexes, reviews, art critiques). The pilot phase of bibliometric analyses was carried out at the Comenius University (CU) in Bratislava on the data of publications register that is available via iPortal of the CU Academic Library (http://alis.uniba.sk:8000/cgi-bin/gw_49_3_8/chameleon?skin=epc).

In October 2013, the CU database of publication outputs, which has been systematically built since the beginning of 1990s, contained 226 453 records of various types of publications and 212 787 citations. The distribution of records among faculties in the whole retrospect of the database is as in Table 1.

Table 1 CU database of publications outputs overview

Faculty [†]	Number of records/publications	Number of citations	Number of cited publications
FMED	50 388	49 873	8 696
JFMED	19 340	22 380	3 698
FPHARM	13 415	12 872	2 574
FLAW	8 991	5 464	1 563
FPHIL	39 311	35 529	8 758
FNS	38 674	72 923	11 379
FMPH	21 022	41 288	4 726
FSPORT	11 094	12 923	2 744
FEDU	17 036	12 334	3 050
FEVTH	2 439	545	244
FRCTH	4 898	2 269	885
FM	3 990	3 242	975
FSEV	2 529	1 250	643
CU Total*	226 453	212 787	34 968

* Total number of publications/citations is not sum of faculties because of co-authorship among authors from different CU faculties.

From the digital repository point of view, we were interested in the typological structure of publications. We focused especially on publications which were published recently - during the period of years 2011-2013 (till October 2013). Within these 3 years CU authors published 25 281 publications (about 9.000 publications per year), in the structure described in Table 2.

When we analyzed given data by faculties, we had noticed differences in preferred publication channels between scientific disciplines (distinction between hard and soft sciences), see Graph 1. It has been proved, in compliance with similar analyses published abroad, that social sciences and humanities' outputs differ from those from natural sciences, medicine and technology. We can observe higher preference of monographic literature (scientific monographs, textbooks, chapters) in soft sciences; hard sciences mostly prefer scientific journals as communication channels. Quite interesting is the distribution of conference papers, which represent 22% of soft sciences publications and only 8% of hard sciences publications.

[†] CU faculties (13): Faculty of Medicine, Jessenius Faculty of Medicine in Martin, Faculty of Pharmacy, Faculty of Law, Faculty of Arts, Faculty of Natural Sciences, Faculty of Mathematics, Physics and Informatics, Faculty of Physical Education and Sport, Faculty of Education, Evangelical Lutheran Theological Faculty, Faculty of Roman Catholic Theology of Cyril and Methodius, Faculty of Management, Faculty of Social and Economic Sciences

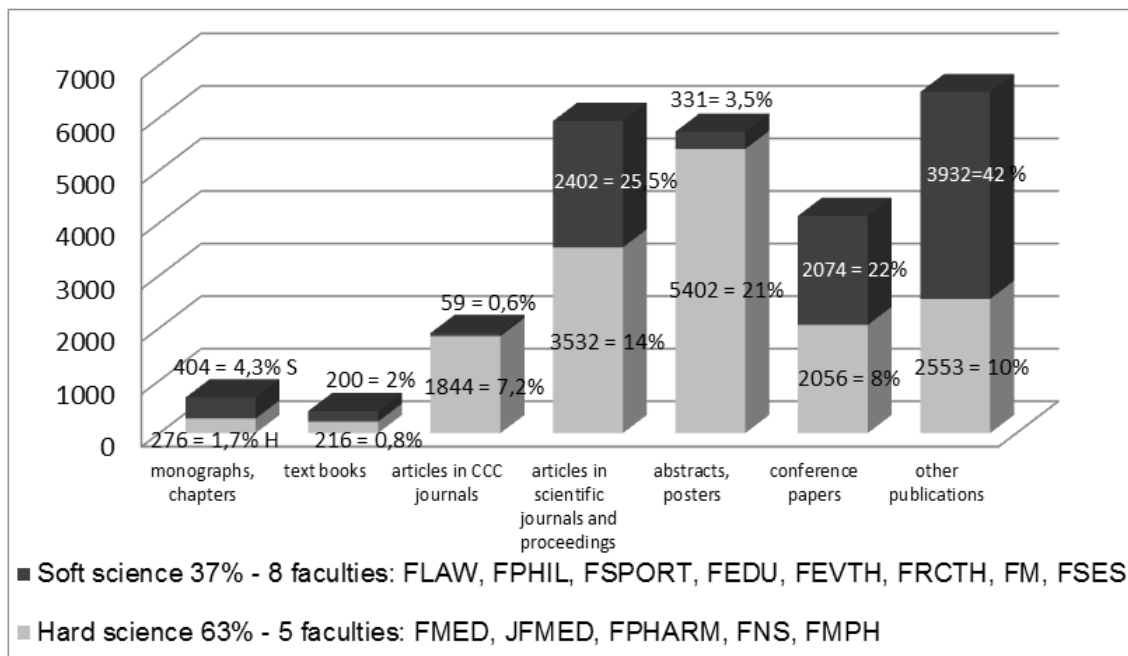
Table 2 CU publications 2011-2013 – typological structure

Typological structure of CU publications	Number of publications	% of all publications
monographs, chapters	688	2.69%
text books	441	1.65%
articles in CCC journals	1 911	7.53%
articles in scientific journals and proceedings	5 958	23.47%
abstracts, posters	5 649	22.68%
conference papers	4 120	16.34%
other publications	6 514	25.65%
Total	25 281	100%

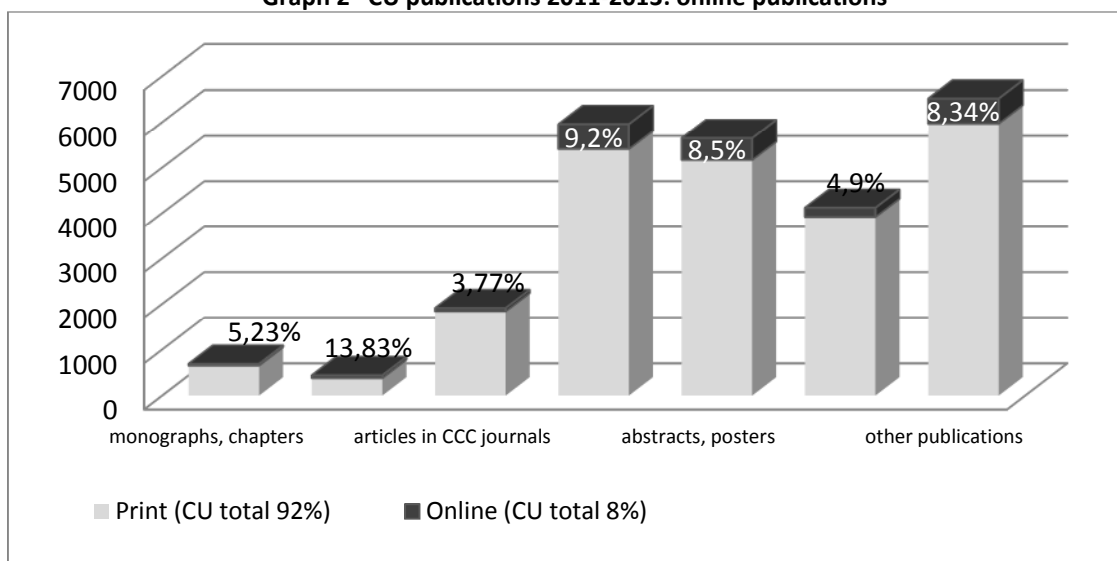
Furthermore, we wanted to know the percentage of publications published online in electronic form, as these could be directly used for repository needs (after solving copyright matters, of course). Only 8% of CU publications from 2011-2013 had been published online (for their distribution among faculties see Graphs 3, 4 and Table 3).

Our analyses showed that Faculty of Mathematics, Physics and Informatics (FMPH), Faculty of Arts (FPHIL) and both medical faculties (FMED, JFMED) belong among faculties with highest rate of online publications. In case of FMPH it is certainly also due to quite long tradition of publishing preprints via well-known specialized web portals such as arxiv.org and cern.ch. Medical faculties publish a lot of abstracts which are often available via different web sites, medical portals and databases (PubMed, bmj.sk). FPHIL online publications are represented especially by domestic conference proceedings, peer-reviewed collections of papers and articles in Slovak online journals.

Graph 1 CU publications 2011-2013: different faculties = different publication channels



Graph 2 CU publications 2011-2013: online publications



Graph 3 CU publications 2011-2013: online publications by faculties

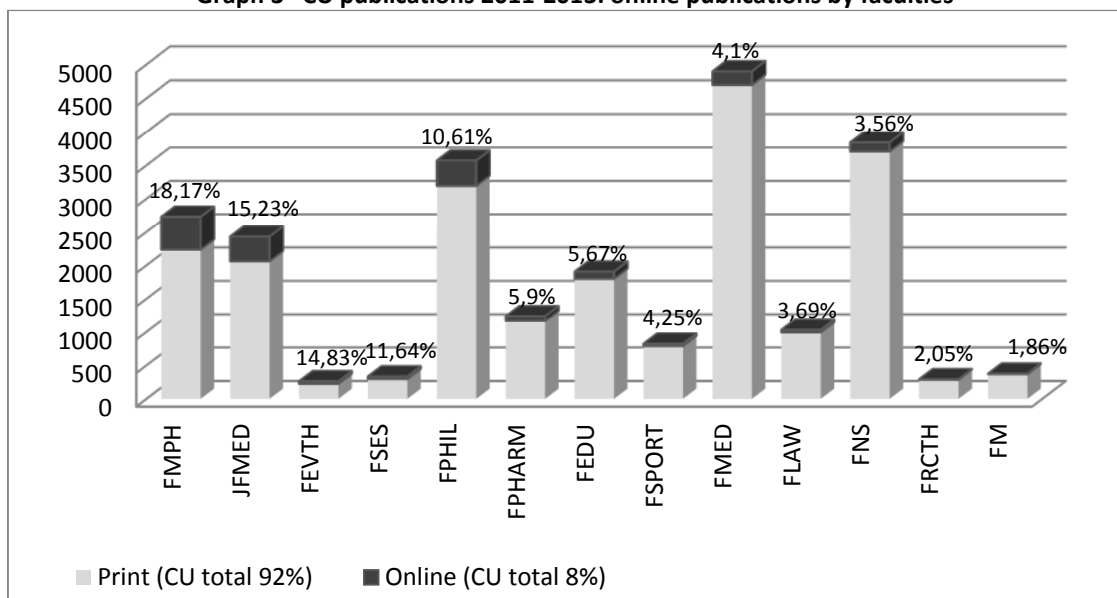


Table 3 CU publications 2011-2013: share of online publications by faculties

CU faculties	Number of online publications	% of all online publications
FMPH	496	25,73%
FPHIL	379	19,66%
JFMED	370	19,19%
FMED	201	10,43%
FNS	137	7,11%
FEDU	108	5,60%
FPHARM	73	3,79%
FSES	39	2,02%
FEVTH	39	2,02%
FLAW	38	1,97%
FSPORT	35	1,82%
FM	7	0,36%
FRCTH	6	0,31%

Almost 17% of all CU online publications are accessible from CU domain (uniba.sk) and its component parts, they are distributed on 48 various locations (e.g. stella.uniba.sk, motiv.fns.uniba.sk, ii.fmph.uniba.sk, moodle.uniba.sk, etc.). In addition, CU operates several local storages with purpose to store and make available different documents in electronic form, especially theses (> 52 400), digitalized collections of papers published by CU (> 400), textbooks (> 50). Such distribution of documents complicates their discoverability as well as preservation. And we managed to prove, that better visibility and discoverability of publications achieved by online publishing channels could increase average citation number. It is obvious especially in following publication categories: scientific domestic conference proceedings and collections of papers; articles in domestic journals, textbooks, articles in CCC journals, articles in foreign scientific journals.

5. Conclusions

Establishing an institutional repository is still quite complicated in Slovak academic institutions, although the benefits of the repository for academic individuals (e.g. increased dissemination and impact of the research outputs, provision of archiving) as well as for the institution itself (effective marketing, better support of education and research, central archives) are well known.

Comenius University in Bratislava feels the need to build complex institutional repository, as up to now it had been using only locally distributed storages enabling access to some of its electronic publications. Institutional repository could be built within CU scientific park and then it could serve as an institutional showcase for research. But effective operation of the repository requires first of all well thought out institutional information strategy, collaboration between researchers/authors, departments, faculties, libraries and other component parts of university. Last but not least it is necessary to deal with copyright and legislation problems.

Acknowledgements

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0208-10, the project of the "University Scientific Park" of Comenius University in Bratislava ITMS 26240220086 and project KEGA 133UK-4/2013 MIKS - Mediálne a informačné kompetencie pre znalostnú spoločnosť.

References

- BJÖRK, Bo-Christer and Ziga TURK. 2000. A Survey of the Impact of the Internet on Scientific Publishing in Construction IT and Construction Management. In: *The Journal of Information Technology in Construction (ITcon)* [online]. Rotterdam: International Council for Research and Innovation in Building and Construction, 2000, vol. 5, pp. 73-88 [ret. 2013-11-04]. ISSN 1874-4753. Available from: <http://www.itcon.org/2000/5/paper.htm>
- BUTLER, Linda. 2008. Using a balanced approach to bibliometrics: quantitative performance measures in the Australian Research Quality Framework. In: *Ethics in Science and Environmental Politic. Ethics Sci Environ Polit* [online]. Oldendorf (Luhe): Inter-Research. June 2008, vol. 8, no.11, pp. 83-92, 2008 [ret. 2013-11-04]. ISSN 1611-8014. Available from: <http://www.int-res.com/articles/esept2008/8/e008p083.pdf>
- COX, Raymond A. K., CHUNG, Kee H. and Albert A. OKUNADE. 1993. Publishing behavior of individuals and most prolific authors in the economics literature. In: *Quarterly Journal of Business and Economics* [online]. Lincoln: University of Nebraska-Lincoln College of Business Administration. June 1993, vol. 32, no. 3, pp. 32-42 [ret. 2013-11-04]. ISSN 1939-8123. Available from: <http://www.allbusiness.com/finance/404806-1.html>
- CRAIG, Iain D., PLUME, Andrew M., MCVEIGH, Marie E., PRINGLE, James and Mayur AMIN. 2007. Do Open Access Articles Have Greater Citation Impact? A critical review of the literature. In: *Journal of Informetrics* [online]. Amsterdam: Elsevier BV. July 2007, vol. 1, iss. 3, pp. 239-248 [ret. 2013-11-04]. ISSN 1751-1577. Available from: <http://dx.doi.org/10.1016/j.joi.2007.04.001>
- ERRAMI, Mounir and Harold GARNER. 2008. A tale of two citations. In: *Nature* [online]. London: Nature Publishing Group, 2008, vol. 451, pp. 397-399 [ret. 2009-08-20]. ISSN 0028-0836. Available from: <http://www.nature.com/nature/journal/v451/n7177/full/451397a.html>
- HARNAD, Stevan and Tim BRODY. 2004. Comparing the impact of Open Access (OA) vs. non-OA articles in the same journals. In: *D-Lib Magazine* [online]. Reston: Corporation for National Research Initiatives. June 2004, vol. 10, no. 6 [ret. 2013-11-04]. ISSN 1082-9873. Available from: <http://www.dlib.org/dlib/june04/harnad/06harnad.html>

HOUGHTON, John and Alma SWAN. 2013. Planting the Green Seeds for a Golden Harvest: Comments and Clarifications on "Going for Gold". In: *D-Lib Magazine* [online]. Reston: Corporation for National Research Initiatives. January/February 2004, vol. 19, no. 1/2 [ret. 2013-11-04]. ISSN 1082-9873. Available from: <http://www.dlib.org/dlib/january13/houghton/01houghton.html>

KIMLIČKA, Štefan. 2001. Význam spracovania a prezentácie publikačnej činnosti na vysokej škole - úloha akademických knižníc. In: *Zborník z odborného seminára Evidencia publikačnej činnosti, citácie a medzinárodné štandardy, ktorý sa uskutočnil 30. januára 2001 v Nitre* [online]. Nitra: Slovenská poľnohospodárska knižnica pri SPU v Nitre. February 2001 [ret. 2013-11-04]. Available from: <http://www.slpk.sk/eldo/epc/1.htm>

KYVIK, Svein. 2003. Changing trends in publishing behavior among university faculty, 1980-2000. In: *Scientometrics* [online]. Budapest: Akadémiai Kiadó Rt. September 2003, vol. 58, iss. 1, pp. 35-48 [ret. 2013-11-04]. ISSN 1588-286. Available from: <http://www.springerlink.com/content/t779563774467186/>

LAWRENCE, Steve. 2001. Free online availability substantially increases a paper's impact. In: *Nature*. May 2001, vol. 411, no. 6837, pp. 521. ISSN 0028-0836.

LUUKKONEN, Terttu. 1992. Is scientists' publishing behavior rewardseeking? In: *Scientometrics* [online]. Budapest: Akadémiai Kiadó Rt. June 1992, vol. 24, no. 2, pp. 297-319 [ret. 2013-11-04]. ISSN 1588-286. Available from: <http://www.springerlink.com/content/r5583r778831t1t4/>

MICHELS, Carolin and Ulrich SCHMOCH. 2013. Impact of bibliometric studies on the publication behaviour of authors of authors. In: *Scientometrics* [online]. Budapest: Akadémiai Kiadó Rt. April 2013 [ret. 2013-11-04]. ISSN 1588-286. Available from: <http://link.springer.com/article/10.1007/s11192-013-1015-7>

Predpis č. 456/2012 Z. z. Vyhláška Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky o centrálnom registri evidencie publikačnej činnosti a centrálnom registri evidencie umeleckej činnosti.

REY, J., MARTÍN, M.-J., PLAZA, L., IBÁÑEZ, J. J. and Isabel MÉNDEZ. 1998. Changes in publishing behavior in response to research policy guidelines. The case of the Spanish Research Council in the field of agronomy. In: *Scientometrics* [online]. Budapest: Akadémiai Kiadó Rt. January 1998, vol. 41, iss. 1-2, pp. 101-111 [ret. 2013-11-04]. ISSN 1588-286. Available from: <http://www.springerlink.com/content/q744436514g14743/>

ROCKWELL, Geoffrey and Lynne SIEMENS. 2000. *The Credibility of Electronic Publishing : Report on Responses to the Questionnaire* [online]. Nanaimo: Malaspina University. 2000 [ret. 2013-11-04]. Available from: <http://web.mala.bc.ca/hssfc/Final/QuestionnaireR.htm>

SKALKA, Ján and Libor VOZÁR. 2008. Centrálny register publikačnej činnosti. In: *ITlib. Informačné technológie a knižnice* [online]. Bratislava: Centrum VTI SR. 2008, no. 2, pp. 20-24 [ret. 2013-11-04]. ISSN 1336-0779. Available from: http://itlib.cvtisr.sk/archiv/2008/2/centralny-register-publikacnej-cinnosti.html?page_id=1279

STEINEROVÁ, Jela and Jaroslav ŠUŠOL. 2005. Library users in human information behavior. In: *Online information review* [online]. Bingley: Emerald Group Publishing Ltd. 2005, vol. 29, no. 2, pp. 139-156 [ret. 2013-11-04]. ISSN 1468-4527. Available from: <http://www.emeraldinsight.com/journals.htm?articleid=1464932&show=html&ini=CNKI>

STEINEROVÁ, Jela, ŠUŠOL, Jaroslav and Mirka GREŠKOVÁ. 2006. Information behavior in relevance judgements. In: *Využívanie informácií v informačnej spoločnosti*. Bratislava: Centrum VTI SR, 2006, pp. 29-40. ISBN 80-85165-92-9.

STEINEROVÁ, Jela (edt.). 2012. *Informačná ekológia akademického informačného prostredia*. Bratislava: Univerzita Komenského, 2012. ISBN 978-80-223-3178.

SWAN, Alma. 2008. *Key Concerns Within The Scholarly Communication Process. Report To The Jisc Scholarly Communications Group* [online]. Truro: Key Perspectives Ltd, March 2008, 65 p. [ret. 2013-11-04]. Available from: <http://www.jisc.ac.uk/media/documents/aboutus/workinggroups/topconcernsreport.doc>

ŠUŠOL, Jaroslav. 2004. Publikačné preferencie autorov vo vzťahu k tradičným a elektronickým zdrojom - situácia v oblasti vedeckej komunikácie. In: *Knižničná a informačná veda 20 = Library and information science 20*. Bratislava: Univerzita Komenského, 2004, pp. 37-57. ISBN 80-223-1993-7.

Zákon č. 183/2000 Z. z. o knižniciach, o doplnení zákona Slovenskej národnej rady č. 27/1987 Zb. o štátnej pamiatkovej starostlivosti a o zmene a doplnení zákona č. 68/1997 Z. z. o Matici slovenskej v znení neskorších predpisov.

Zákon č. 455/2012 Z.z. Zákon, ktorým sa mení a dopĺňa zákon č. 131/2002 Z. z. o vysokých školách a o zmene a doplnení niektorých zákonov v znení neskorších predpisov, a ktorým sa mení zákon č. 455/2004 Z. z. o zriadení Akadémie ozbrojených síl generála Milana Rastislava Štefánika, o zlúčení Vojenskej leteckej akadémie generála Milana Rastislava Štefánika v Košiciach s Technickou univerzitou v Košiciach, o zriadení Národnej akadémie obrany maršala Andreja Hadika a o zmene a doplnení niektorých zákonov v znení zákona č. 144/2008 Z. z.



Ila
Information
International
Associates

Navigators in a Sea of Data

For 23 years, Ila has provided quality information services to advance our clients' mission. By applying information science skills in combination with information technology, we help organizations handle massive amounts of information and make it useable.

**Ila makes information an asset
for your organization.**

Corporate Office:
104 Union Valley Road
Oak Ridge, TN 37830
Phone: 865- 481-0388

D.C. Area Office:
417 West Broad Street
Suite 201
Falls Church, VA 22046
Phone: (703) 407-5770



www.iiaweb.com

Back to Grey: Disclosure and Concealment of Electronic Theses and Dissertations*

Joachim Schöpfel and H  l  ne Prost (France)

Abstract

The open access principle requires that scientific information be made widely and readily available to society. Defined in 2003 as a "comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community", open access implies that content be openly accessible and this needs the active commitment of each and every individual producer of scientific knowledge. Yet, in spite of the growing success of the open access initiative, a significant part of scientific and technical information remains unavailable on the web or circulates with restrictions. Even in institutional repositories (IRs) created to provide access to the scientific output of an academic institution, more or less important sectors of the scientific production are missing. This is because of lack of awareness, embargo, deposit of metadata without full text, confidential content etc. This problem concerns in particular electronic theses and dissertations (ETDs) that are disseminated with different status – some are freely available, others are under embargo, confidential, restricted to campus access (encrypted or not) or not available at all. While other papers may be available through alternative channels (journals, monographs etc.), ETDs most often are not. Our paper describes a new and unexpected effect of the development of digital libraries and open access, as a paradoxical practice of hiding information from the scientific community and society, while partly sharing it with a restricted population (campus). The study builds on a review of recent papers on ETDs in IRs and evaluates the availability of ETDs in a small panel of European and American academic IRs and networks. It provides empirical evidence on the reality of restricted access and proposes a model of independent variables affecting decisions on embargo and on-campus access, together with a table of different degrees of (non) open access to ETDs in IRs.

The context

Scientific grey literature stands for intellectual works not controlled by commercial publishers, of sufficient quality to be collected and preserved, but often difficult to obtain. The difficulty of acquisition and collection building was one of the main characteristics of grey literature in the past. The Web changed the situation. Dissemination of scientific information and access to the full text of all kinds of documents became easy. Concerning grey literature, the Web was considered as a solution and at the same time, as the final destination. The idea was simple and convincing: increasing availability and accessibility would change the nature of grey literature and, in the end, make it disappear. Grey would turn into white (Artus 2003). This belief was strongly supported by the success of the movement towards open access to scientific information (Suber 2012). The open access principle requires that scientific information be made widely and readily available to society (Willinsky 2005). Defined in 2003 as a "comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community"ⁱ, open access implies that content be openly accessible and this needs the active commitment of each and every individual producer of scientific knowledge.

The reality is different. Not only the definition of grey literature can (and will) survive the Web and open access, (Schöpfel 2010) but also contrary to all expectations and hopes, the Web sometimes increases barriers to scientific information. In spite of the growing success of the open access initiative, a significant part of scientific and technical information remains unavailable on the web or circulates with restrictions. Even in institutional repositories created to provide access to the scientific output of academic organizations, more or less important sectors of the scientific production are missing. The reasons are multiple: lack of awareness, embargoes, deposit of metadata without full text, confidential content and privacy concerns etc. This problem concerns in particular electronic theses and dissertations (ETDs). Many are freely available, but others are under embargo or confidential, restricted to campus access (encrypted or not) or not available at all. While some professionals and scholars are increasingly concerned about the situation (Owen et al. 2009), others welcome the protection of copyright (Hawkins et al. 2013).

Our paper provides empirical evidence on restricted access to American and European ETDs, reviews some published explanations, and then makes a proposal of a conceptual model of independent

* First published in the GL15 Conference Proceedings, March 2014.

variables affecting decisions on embargo and on-campus access, together with a table of different degrees of (non) open access to ETDs in institutional repositories (IRs).

The paper builds on a study conducted in Lille between January and April 2013 (Schöpfel & Prost 2013) and contributes to a French-German survey on ETD embargoes carried out by the Institute for Science Networking at the University of Oldenburg and the University of Lille 3.

Evidence

A small but growing number of empirical studies on ETDs reveal figures on access restriction. A survey conducted in winter 2013 produced complementary figures from France, Europe and the United States. Table 1 presents figures from fifteen institutions and service providers, with the surveyed number of theses, the percentage of documents without access restriction, and the part of documents under embargo or restricted to campus-only access.

Institution	Sample	% open access	% restricted access	% on campus	% embargo
ProQuest (US)	500000+	95%	5%	n/a	5%
Texas (Austin)	11539	92%	8%	8%	n/a
PUC Rio de Janeiro (Brazil)	2787	89%	11%	n/a	11%
West Virginia	4600+	85%	15%	15%	0.3%
Florida State	3709	84%	16%	n/a	n/a
ABES (France)	10631	80%	20%	20%	n/a
Lille 1 (France)	833	79%	21%	15,5%	5,5%
Auckland (Australia)	3088	72%	28%	28%	n/a
Lorraine (France)	52	71%	29%	29%	n/a
Maryland	2050	68%	32%	n/a	32%
Valenciennes (France)	35	63%	37%	31%	6%
Liège (Belgium)	191	57%	43%	33%	10%
Virginia Tech	20386	54%	46%	n/a	46%
Amherst	n/a	48%	52%	32%	20%
Lille 3 (France)	124	40%	60%	n/a	n/a

Table 1. Empirical evidence on restricted access to electronic theses and dissertations (ETDs)

Taken together, about 10% of these roughly 550,000 electronic theses are not freely available on the Internet. Without the ProQuest figures, this part with limited access rises to 26%, ranging from 10% to more than 50%. 17% are embargoed for six months to two years or longer while the other 9% can only be accessed on-campus. This panel may not be representative and the results should be interpreted with caution. Nevertheless, they point out that the problem is not limited to a country or region but concerns all institutions with ETDs infrastructures and IRs. Some examplesⁱⁱ:

At Amherst College, Massachusetts, 32% of PhD theses cannot be accessed from outside of the campus and 20% are under embargo for at least six months (Banach 2011).

At the University of Maryland, 68% ETDs are available without any restrictions. The other theses are under embargo, 21% up to one year and 11% from one to six years (Owen et al. 2009).

ProQuest Dissertation Publishing conducted in 2012 a study on ten years embargo trends (2000-2010) in the ProQuest Dissertations and Theses (PQDT) database. The surveyed corpus of 500,000+ print and electronic theses contained about 25,000 embargoed items (5%). Most of the embargoes are short-term embargoes, for six months to five years, but a small part of theses are under permanent (long term) embargo.

In Brazil, Pavani & Mazzeto (2009) describe access restriction for 11% ETDs on the campus of the Pontificia Universidade Católica at Rio de Janeiro. About 21% of these files are under embargo for five years or longer.

The University of Liège (Belgium) document server indicates 191 PhD theses for 2012. 108 are freely available on their IR called ORBi (57%). For 33%, the access is limited to the campus; the remaining 10% are embargoed for a non-specified delay.

Since 2006, French universities have progressively switched from the traditional handling of print PhD theses to the new infrastructure of ETDs called STAR, linked to a national gateway "Theses.fr" run by ABES at Montpellierⁱⁱⁱ. From 2006 to 2012, the STAR system processed 10,631 ETDs. 8,737 theses were available on the web without any restrictions (80%) while access to the other 1,894 theses was limited to on-campus availability (20%). STAR does not provide information about embargoes.

Another example from France: from 2008 to 2011, the University of Lille 1 processed 833 ETDs in Science and Technology. Nearly 80% are in open access on their IR. 15% are available on the campus only while the other 5% are under unlimited embargo, based on a decision of the faculty to protect intellectual property and innovation.

Only few data on long-term trends have been published. Based on figures from ProQuest, Hawkins et al. (2013) identified an increasing number of embargoed ETDs. The findings by Owen et al. (2009) can be interpreted in the same way, especially for short-term one-year embargoes. On the other hand, the embargo statistics at West Virginia appear to be relatively stable over time (Hagen 2010), just like the figures between 2008 and 2011 from Lille.

Explanations

Following our review and survey data, experts and professionals explain the access restrictions in different ways, with arguments based on statistics, experience and anecdotic evidence. In a UK survey on mandates for ETDs, 88% of the universities indicated that they allow authors of theses to impose restrictions on access to their work, i.e. the electronic file, for many different reasons. Students, with the agreement of their supervisor, can request an embargo for the following reasons: commercial contract (for instance, funding by an external organisation), patent pending, ethical confidentiality and/or sensitive material (data protection), publication pending and third party copyright (Brown et al. 2010). The same study reveals that restrictions on grounds of third party copyright, data protection or potential risks to personal safety were reported only amongst ETDs (not print support) and that only 60% of the universities allow students to impose restrictions for print theses.

At Brunel University, "while every effort has been made to ensure that embargoing access to theses is not used as 'a panacea against all ills', students are offered the option of a 3-year embargo if they have a publication or patent pending" (Brown & Sadler 2010). Academics of the University of Maryland mention future publication, protection of data or work, student request, proprietary data and patent application as primary reasons for approving of embargoes (Owen et al. 2009).

In France, PhD theses are considered as administrative documents and (except confidential research) must be disseminated, at least on the campus (Schöpfel & Lipinski 2012). Yet, following our survey at Valenciennes (France) PhD students appear sometimes confused by embargo, confidentiality and on-campus options.

In Italy, Arabito et al. (2008) justify embargo options as indispensable for the same reasons: "(...) the free availability of doctoral theses on the web can be jeopardized by thorny copyright issues, which arise in the following cases: use of third party owned materials (...), third parties involved (possible infringement of privacy), patentable discoveries (...), and ongoing publication of data (according to the publisher policy)".

This last argument – expected publication – is by far the most common reason and explains between 1/3 (Owen et al. 2009) and 3/4 (Pavani & Mazzeto 2009) of all embargo decisions. The role of faculty appears to be crucial. At Virginia Tech, nearly half of the students' embargo decisions were taken on advice by faculty while requests by publishers are insignificant (McMillan et al. 2012). Ramirez et al. (2013) confirm that "scholars continue to doubt the viability of publishing opportunities after a dissertation or thesis becomes available electronically in an open access repository. Perceptions and fear, not data, inform many graduate advisors' and graduate students' decisions to restrict access to their ETDs".

Each graduate school has its own guidelines. A recent survey with more than 150 American graduate schools show that nearly 30% of all institutions "either don't allow an embargo at all, or don't tell students (about it at all) where they can find that information readily (...). In their enthusiasm for OA, universities and libraries across the U.S. are cajoling, arm-twisting, or even coercing students into in effect surrendering the copyright to their dissertations and theses, sometimes with the threat that students cannot graduate if they disagree" (Hawkins et al. 2013).

Florida State University Graduate School implemented access restriction - on campus only access – for older, digitized PhD theses: "Since retrospective digitized theses and dissertations did not include retrospective digitized access agreement forms, senior leadership recommended IP restriction for all FSU retrospective digitized theses and dissertations in 2009" (Smith 2009).

Kleister et al. (2013) report how changing the embargo policy at the University of North Texas dramatically dropped down the number of embargoed ETDs, from 80-100 to 20 or less per year. Asking for embargo has always been possible but the burden was on the PhD student to initiate the discussion. From the moment (2007) when this "burden" was replaced by a simple option on the agreement form (as check boxes), the number of embargo decisions was multiplied by more than five. Their conclusion is clear: "The needs of students must be balanced against the institution's needs and goals. Justification for embargo should not be especially onerous, but needs to be more than a mere checkbox on a form..."

At West Virginia University, Hagen (2010) reports that for the period 1998-2010, 85% of the more than 4600 theses are disseminated without any restriction. The part of theses with restricted access decreased from 47% (1998-2000) to 15% in 2010, because the option of encrypted on-campus only access was phased out in 2009 while the part of embargoed ETDs remained stable.

Smith (2009) describes how the Florida State University Graduate School requested campus-community and PDF document security options starting in Fall 2008, and he adds that “since retrospective digitized theses and dissertations did not include retrospective digitized access agreement forms, senior leadership recommended IP restriction for all FSU retrospective digitized theses and dissertations in 2009”. Following the published figures, this part of restricted access can be estimated at about 16%. Only three studies present detailed embargo statistics cut down by scientific disciplines (Owen et al. 2009, Pavani & Mazzeto 2009, ProQuest 2012). Yet, these survey results are not really reliable. Some disciplines appear to be relatively consistent, such as life and chemical sciences, agriculture and environment, business, some domains of engineering (applied sciences) and public health, all with medium or high rates of embargoes. Pavani & Mazzeto (2009) show that in Science and Technology, pending publications as a reason for embargo concern mostly articles (73%) while in Social Sciences students intend above all publishing a book (57%). Yet, we must be careful with these statistics because of more or less small samples.

People, institutions, reasons and objectives

At first glance, the situation appears rather simple. PhD theses being intellectual work, the student is the only person holding the right to decide about dissemination. Of course, this view is by far much too simplistic. Different actors – people and institutions – can be distinguished who impact more or less the process of decision-making, with different reasons, motivations and objectives. A non-exhaustive list may be helpful to distinguish the different participants in this decision-making process:

- *PhD student: may want to keep the rights to his/her intellectual work; receives advice or orders from the different actors of his/her scientific community*
- *Director of PhD thesis: concerned by quality and reputation, fear of plagiarism.*
- *Jury: concerned by quality and reputation, protection of results.*
- *Community (discipline, staff): supportive or indifferent attitudes towards open access.*
- *Other PhD students: shared concerns about career, evaluation, and plagiarism...*
- *Graduate school: favourable or indifferent towards open access.*
- *University presidency (dean): supporting or not open access policy; concerned with third party rights (confidentiality, copyright infringement).*
- *Academic library: often in favour of open access and running an institutional repository.*
- *Service provider: supportive or indifferent towards open access.*
- *Publishers: opposed or not to open access and publishing of OA theses.*

Figure 1 tries to map these players in a system of decision-making of dissemination and access to ETDs. Each player sets his own goals, fulfils specific functions, plays his particular role, sometimes consistent with others, sometimes in opposition. For instance, PhD students may deposit their non-reviewed papers in open archives “off-campus”, outside of their institution and without any validation or authorization, even when the jury rejects the disclosure.

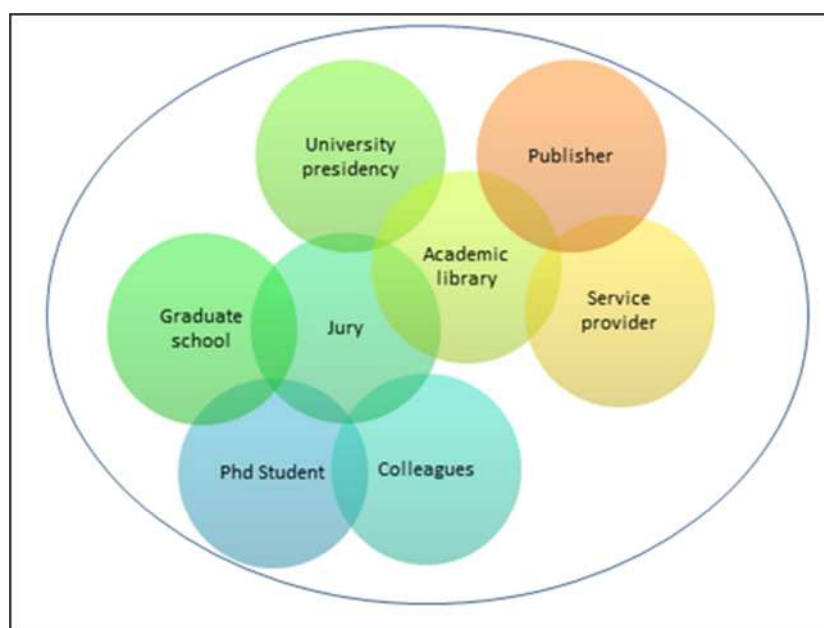


Figure 1: Participants in the decision-making on dissemination of ETDs

All these people, groups and institutions act in different ways, for different reasons, with different objectives and strategies. The literature review and survey results reveal the following components that may be understood as independent variables of the final decision:

- *A publishing project (article, book): if the student intends to publish his results with a scientific publishing house, he/she may be reluctant to disseminate the thesis on the Internet.*
- *Individual knowledge (or ignorance) of publishers' policies towards publishing papers that are already available on the Internet.*
- *Individual attitude toward open access (awareness, ethics, risk avoidance).*
- *Institutional decision on confidentiality and dissemination.*
- *Legal environment (copyright, intellectual property, disclosure of PhD theses).*
- *Institutional open access policy (awareness, risk avoidance).*
- *Institutional workflow of processing ETDs (reference points, opt-ins or opt-outs, easiness).*
- *Protection of third parties' rights (intellectual property, confidentiality, privacy).*
- *The jury members' advice (quality and excellence, awareness of open access).*
- *Tradition and attitudes of the scientific community.*
- *Publishers' acceptance of open access papers: If the publisher does not accept papers or books based on theses openly available on the Internet, his attitude may foster decisions in favor of embargoes.*

Each of these aspects acts in a different way. Some elements may decide on dissemination or non-dissemination, while others are limited to embargo or on/off-campus decisions. Moreover, some are case-by-case decisions while others reflect general attitudes and stable behaviours. Again, a schema may be helpful for global understanding (figure 2):

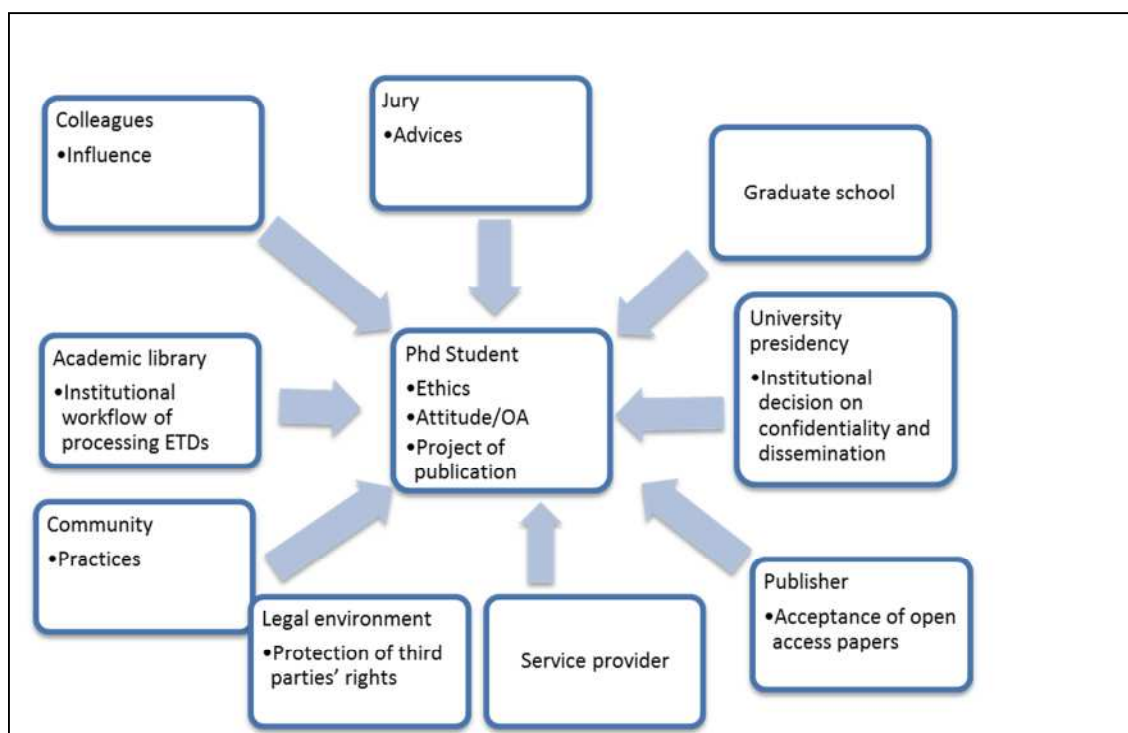


Figure 2: Different aspects of decision-making on dissemination and concealment of ETDs

This model may need empirical confirmation and perhaps, more details. Yet, its central characteristics are the multi-factorial or multivariate approach to the prediction of decisions on dissemination or concealment of ETDs. Even if individual publishing strategies and attitudes towards open access may play a major role, other variables such as personal advice from the PhD director, easiness of decision and references should not be neglected, in particular when discussing ways of improving accessibility and availability of PhD theses.

A model of openness

With regards to accessibility and availability of PhD theses, our analysis showed so far that openness is not a simple, binary concept but that the documents can be more or less open, depending on different variables. Some of those variables are similar to articles published in journals or books, but others are specific to PhD theses. In October 2012, the *Scholarly Publishing and Academic Resources Coalition* (SPARC), “an international alliance of academic and research libraries working to create a more open system of scholarly communication”^{iv}, released a guide called “How Open Is It” that outlines the core components of open access (e.g., reader rights, reuse rights, copyrights, author posting rights, etc.) across the continuum from “open access” to “restricted access”. Compared with our multivariate approach, this *Open Access Spectrum* (SPARC 2012) helps to get a realistic view on the problems of openness, disclosure and concealment of theses. Table 2 shows a possible adaptation of the SPARC guide to the specific conditions of the dissemination of PhD theses.

Reader rights: On site only, or also at distance, via authentication? What about interlibrary loan or document delivery?

Reuse rights: Generous reuse rights (CC-BY licensing) or full copyright protection?

Copyrights: No third party claim or complete concealment (confidentiality) because of sensible or protected results?

Institution rights: In France, by decree, the PhD theses must be disclosed, at least on the campus, except for confidential projects.

Institution policy: In fact, at least two different levels must be distinguished, the global policy of the university (or faculty/department), and the approach of the jury which may, in some cases and sometimes for political reasons, reject open access disclosure via institutional repository even for non-confidential theses.

Posting workflow: Following empirical studies on changes in ETD workflows, we adapt the SPARC component “Automatic posting” to the specific ETD environment. The continuum between open and closed ranges from procedures without embargo options, i.e. where an embargo decision needs a specific individual action (written and argued request), to workflows where open access is available only as an opt-in option while on-campus dissemination is the default option.

Machine readability: The last component is about automatic access and exploitation of the full text and the related data and metadata. Exploitation means: text or data mining, harvesting, or crawling. Our table summarizes the SPARC scale but specifies the existence of supplementary data files (tables, videos, images etc.) that may have been submitted together with the thesis.

Table 2: Degrees of openness for electronic theses in institutional repositories (adapted from SPARC 2012)

Access	Reader Rights	Reuse Rights	Copyrights	Institution Rights	Institution Policy	Posting Workflow	Machine Readability
Open	Free readership rights immediately upon posting	Generous reuse & remixing rights (CC-BY license)	No third party claim	Obligation of free dissemination	OA mandat	No embargo option available in workflow, embargo needs special procedure	Full text, metadata, citations & data, incl supplementary data (annex etc), provided in machine-readable standard formats through a community standard API or protocol
	Free readership rights after 6 months embargo	Reuse, remixing & further building upon the work subject to certain restrictions & conditions (CC BY-NC & other CC licenses)			Institutional repository		Full text, metadata, citations & data, incl supplementary data (annex etc), may be crawled or accessed through a community standard API or protocol
	Free readership rights after 12+ months embargo	Reuse (no remixing or further building upon the work) subject to certain restrictions and conditions	Third party requires embargo	Obligation of dissemination with restrictions (campus)	OA support	Opt-out if claim for embargo	Full text, metadata & citations may be crawled without special permission or registration
	Free readership rights only on-campus (Intranet)				Interest for OA		Full text, metadata & citations may be crawled with permission
Closed	Total embargo (confidentiality)	No reuse rights beyond fair use/limitations & exceptions to copyright (all rights reserved copyright)	Third party claims full IP protection (confidentiality, privacy)	No institution rights	No OA policy	Opt-in for free dissemination	Full text & metadata not available in machine-readable format

Back to grey?

A good idea does not necessarily guarantee success. Internet is not synonymous with openness, and the creation of institutional repositories and ETD workflows does not make all items more accessible and available. Sometimes, the new infrastructure even appears to increase barriers to PhD theses.

Different reasons contribute to this unexpected (and most often, unwanted) development, and in a certain way, new technologies and digital infrastructures trigger the tendency for access restrictions. In our first paper, we discussed empirical data in terms of ethics, law, legitimate interests and policy, trade secrets, individual and institutional strategies and workflow-biased decision-making. Our present communication adds a conceptual framework and a differential description of the specific conditions of this part of scientific communication.

Open access is without doubt a valuable and important goal for scientific communication. Yet, scientific and technical information, considered as a part of research behaviour and object of strategic decisions (Roosendaal et al., 2010) always included decisions on concealment and parts of secrecy. Together with copyright and technologies, these individual and institutional decisions contribute today to an unsatisfying and inefficient situation where one part of digital PhD theses are easy to find and to obtain while others remain hidden, embargoed and/or limited to on-campus access. As for open access and institutional repositories in general, one part of the research community is (so far) indifferent or hostile to unprotected dissemination of theses. From the moment the decision on dissemination of ETDs moves from institution to the individual author, we have to deal with these attitudes and opinions.

Openness is not enough for scientific communication. Internet does not change grey literature into white in a mechanical way. Without a minimum of quality and standardization (Dobratz & Scholze 2006), without metadata, referencing, long-term preservation, discovery tools etc., (Schöpfel et al. 2011), and without raising awareness, thorough decision aids and redesigned workflows, perhaps even changes in the legal status of theses, institutional repositories will only provide a partial answer to the question of grey literature in the digital environment. So, back to grey?

Bibliography

- S. Arabito, et al. (2008). 'Time to Harvest: Electronic Doctoral Theses in Italy'. In *ETD 2008 11th International Symposium on Electronic Theses and Dissertations*, 4 - 7 June 2008, The Robert Gordon University, Aberdeen, UK.
- H. M. Artus (2003). 'Old WWWine in New Bottles? Developments in electronic information and communication: structural change and functional inertia'. In *Fifth International Conference on Grey Literature: Grey Matters in the World of Networked Information*, Amsterdam, 4-5 December 2003.
- M. Banach (2011). 'The Benefits of Managing and Publishing ETDs "In House" Using an Open Access Repository'. In *USETDA 2011: The Magic of ETDs...Where Creative Minds Meet. May 18-20, Orlando, Florida*.
- S. Bok (1984). *Secrets: On the Ethics of Concealment and Revelation*. Oxford University Press., Oxford.
- J. Brown & K. Sadler (2010). 'Vision, impact, success: mandating electronic theses. Case studies of e-theses mandates in practice in the UK Higher Education sector'. Tech. rep., UCL University College of London.
- J. Brown, et al. (2010). 'Influencing the Deposit of Electronic Theses in UK HE: report on a sector-wide survey into thesis deposit and open access'. Tech. rep., UCL University College of London, London.
- S. Dobratz & F. Scholze (2006). 'DINI institutional repository certification and beyond'. *Library Hi Tech* 24(4):583–594.
- J. H. Hagen (2010). 'Managing ETDs and Evolution of the Repository: Twelve Years of Case Studies at West Virginia University'. In *ETD 2010 13th International Symposium on Electronic Theses and Dissertations. June 16-18, 2010, Austin, Texas*.
- J. Kleister, et al. (2013). 'Riding the Restriction Wave at the University of North Texas: Comparing IR and ProQuest Embargoes'. In *Third Annual United States Electronic Thesis and Dissertations Association (USETDA) Conference, 2013, Claremont, California, United States*.
- G. McMillan, et al. (2012). 'Do Open Access ETDs Affect Publishing Opportunities in the Sciences?'. In *CNI Fall Membership Meeting, December 12, 2012*.
- T. M. Owen, et al. (2009). 'ETDs in Lock-Down: Trends, Analyses and Faculty Perspectives on ETD Embargoes'. In *ETD 2009 12th International Symposium on Electronic Theses and Dissertations. June 10-13, 2009, Pittsburgh, Pennsylvania*.
- A. Pavani & A. Mazzeto (2009). 'Restricted ETDs and open access'. In *ETD 2009 12th International Symposium on Electronic Theses and Dissertations. June 10-13, 2009, Pittsburgh, Pennsylvania*.
- ProQuest (2012). 'Ten Year Embargo Trends'. In *CGS Summer Workshop and New Deans Institute, Boston, Massachusetts, July 7 – July 11, 2012*.
- M. L. Ramirez, et al. (2013). 'Do Open Access Electronic Theses and Dissertations Diminish Publishing Opportunities in the Social Sciences and Humanities? Findings from a 2011 Survey of Academic Publishers'. *College & Research Libraries* 74(4):368–380.
- H. E. Roosendaal, et al. (2010). *Scientific Publishing: From Vanity to Strategy*. Chandos Publishing, Oxford.
- J. Schöpfel & T. A. Lipinski (2012). 'Legal Aspects of Grey Literature'. *The Grey Journal* 8(3):137–153.
- J. Schöpfel & H. Prost (2013). 'Degrees of Secrecy in an Open Environment. The Case of Electronic Theses and Dissertations'. *ESSACHESS - Journal for Communication Studies* 6(2 (12)).
- J. Schöpfel (2010). 'Towards a Prague Definition of Grey Literature'. In *Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues. Prague, 6-7 December 2010*, pp. 11–26, Amsterdam. Text Release.
- J. Schöpfel, et al. (2011). 'Open is not enough. A case study on grey literature in an OAI environment'. In *Thirteenth International Conference on Grey Literature: The Grey Circuit. From Social Networking to Wealth Creation. Washington, DC, 5-6 December 2011*, pp. 75–86, Amsterdam. Text Release.
- P. Smith (2009). 'ETDs, IRs, Collaborations, and Access at Florida State University'. In *ETD 2009 12th International Symposium on Electronic Theses and Dissertations. June 10-13, 2009, Pittsburgh, Pennsylvania*.
- SPARC, PLOS, OASPA (2012). 'How Open Is It? Open Access Spectrum'.
- P. Suber (2012). *Open access*. MIT Press, Cambridge MA.
- J. Willinsky (2005). *The Access Principle: The Case for Open Access to Research and Scholarship (Digital Libraries and Electronic Publishing)*. MIT Press, Cambridge MA.

All websites accessed in September 2013.

ⁱ Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities of 22 October 2003 <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung/>

ⁱⁱ For more details and examples, see Schöpfel & Prost (2013).

ⁱⁱⁱ Gateway to French theses at <http://www.theses.fr> run by the French Bibliographic Agency of Higher Education ABES <http://www.abes.fr>

^{iv} <http://www.sparc.arl.org/about>

On the News Front

Investigation into the use of Grey Literature in Evidence-Based Medicine: The role of the NHS Library and Information Services Professional

By Joanna Hooper, United Kingdom

Background

Whilst the official definition of Grey Literature(GL) is “That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers” (Farace, 2010), GL has also been described as ephemeral, fugitive and “narrative and descriptive information” (Korjonen, 2013) and it is this perception of GL which has led to it not always sitting comfortably within the field of Evidence Based Medicine(EBM). EBM as a discipline arose from the need to ensure that health practitioners followed clear processes to guarantee predictable outcomes. Sackett et al demonstrate that the origins of EBM grew from the need to ensure that clinicians employed “the conscientious, explicit and judicious use of current best evidence in making decisions about care of individual patients” (Sackett, 1996). EBM depends on standardised, quantitative, measurable data, usually published through the domain of the peer review process, in order to establish the proven reproducibility, and therefore reliability, of data.

It is important to note that EBM does not come without its own problems. Within the scope of EBM, in effect, all literature is considered grey before it has been published in a peer-reviewed journal, and it has been argued that the processes that have been put in place to ensure systematic and quick publication of evidence are not entirely reliable. Much of the peer-review process is opaque and the close involvement of commercial interests and pressures of time result in skewing which clinical trials and studies are chosen to be peer-reviewed and, ultimately, published. (Gura, 2002; Poschl, 2010).

What this means is that trials and studies that are, for example, long and costly to duplicate often remain in the field of GL with the result that valuable data often remains concealed and unaccounted for and this inevitably leads to concerns of bias. McAuley states that “The exclusion of grey literature from meta-analyses can lead to exaggerated estimates of intervention effectiveness.” (McAuley, 2002). Hopewell asserts that “people conducting systematic reviews...need to ensure they search for trials in the grey, as well as published, literature in order to minimise bias” (Hopewell, 2007). Notably Hopewell also maintains that “there is limited evidence to show whether grey trials are of poorer methodological quality than published trials” (Hopewell 2007)

Such studies have established that GL is a necessary element in the field of EBM however, with the development of desk top publishing and the rise of the internet there has been a substantial increase in the amount of information being created, which means that not only is more GL being produced but also, because of these new production techniques, its presence is more visible than it once was. (Farace, 2010). This being said GL still encumbers both the researcher and the Library and Information Services (LIS) professional with many of the same problems that it did in the past, such as ease of access and bibliographic reliability.

Whilst progress has been made by the international grey literature community to promote procedures to standardise production and stabilise bibliographic control it is fair to say that some of the challenges associated with GL look set to remain with us for near future at least, and although Farace and Schopfel state that “New tools for collecting, depositing and archiving will make grey literature less ephemeral and volatile than in the past” they also acknowledge that “bibliographic control of grey literature will remain problematic...” (Farace, 2010). Historically, it is concerns about the reliability, and specifically the reproducibility, of data that has impeded the use of GL in EBM.

Aims and objectives

It is the repercussions of this still very awkward nature of the use GL in EBM, and the impact that this has on NHS Librarians in providing support and teaching searching techniques in accessing evidence in healthcare, that is the focus of this research. The aims and objectives are to discover what NHS Librarians already know about GL and what their perceptions, opinions are and attitudes are towards GL. What are their opinions with regard to some of GL's perceived characteristics, such as access and reliability? Do they set out specifically to search for grey literature when supporting medical and clinical staff in their research? Have NHS Librarians developed any particular search strategies for finding GL? Have they encountered barriers to accessing GL? And if so what are they? It is hoped that through establishing current practice, opinions and attitudes of NHS Librarians that the grey literature community can gain a better insight the perception of the value and use of GL and will hope to identify ways in which healthcare librarians can be supported in further developing their use and recognition of GL as a valuable aspect of EBM.

Methodology

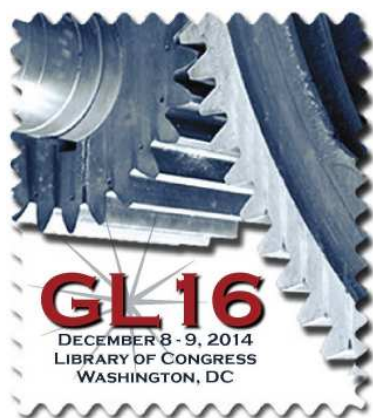
A questionnaire has been developed using an online survey tool, based on a mixed methods approach which consists of questions designed to measure attitude on a scale such as Likert's followed by an open text comment box asking the respondent to give a more reflective answer in order to expand on their reasons for choosing to answer in the way that they have. The questionnaire has been sent to all NHS England regional library leads who will then disseminate it across their regions. It is expected that in this way the questionnaire will reach NHS staff that are involved in supporting NHS healthcare staff with searching for and teaching how to search for evidence in medicine.

Jo Hooper has worked at University Hospitals Bristol NHS Foundation Trust as both a Library Assistant and Librarian. She holds a BA Hons Social Sciences from the University of the West of England (UWE) and is currently studying for the MSc Information and Library management also at UWE. With experience of delivering workshops to NHS staff introducing Grey Literature as an important aspect of clinical research in Evidence-based Medicine, her MSc research focuses on how NHS librarians perceive GL, what strategies they have in place for searching for and delivering GL to NHS staff and whether they promote its use as an aspect of their teaching sessions aimed at Finding the Evidence. Other interests include Open Access and Open Peer Review. Email: jo_hooper@rocketmail.com

References

- Farace, Dominic J. and Schöpfel, Joachim (Eds.). (2010). Grey literature in library and information studies. Berlin: De Gruyter Saur.
- Gura, T., (2002). Scientific publishing: Peer review, unmasked. *Nature*, 419(6878) p.258-260
- Hopewell, S., (2007). Grey Literature in Meta-analysis of randomized trials of health care interventions (Review). *Cochrane Database of Systematic Reviews*. [Online] Available at: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.MR000010.pub3/pdf> [Accessed 15/3/2014]
- Korjonen, H., (2013). *Grey literature in public health: valuable evidence?* [Online] Available at: http://nhfshare.heartforum.org.uk/RMAssets/NHFReports/Grey_literature.pdf [Accessed 28/3/2014]
- McAuley, L., (2002). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses?. *The Lancet*, 356(9237) p.1228-31
- Poschl, U., (2010). Interactive Open Access Publishing and Peer Review: The Effectiveness and Perspectives of Transparency and Self-Regulation in Scientific Communication and Evaluation. *Liber Quarterly: The Journal of European Research Libraries*, 19(3/4), [Online] Available at: http://www.atmospheric-chemistry-and-physics.net/pr_acp_poschl_liber_quarterly_2010_interactive_open_access_publishing.pdf [Accessed 9/12/2013]
- Sackett, D. L., (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, 312:71. 312(71). [Online] Available at: <http://www.bmj.com/content/312/7023/71> [Accessed 15/3/2014]

Sixteenth International Conference on Grey Literature



Grey Literature Lobby: Engines and Requesters for Change

The Library of Congress, Washington DC, USA

Federal Library Information Network, December 8-9, 2014

Call for Posters

- Think Tanks • Advocacy Groups • Research Teams • NGOs • IGOs • Foundations
- Coalitions • Lobbyists • Special Interest Groups • Task forces • Associations, Etc.

Guidelines for Abstracts

Title of Poster:

Topic/Theme:

Author Name(s):

Telephone:

Organization(s):

Fax:

Address/P.O. Box:

Email:

Postal Code - City - Country:

URL:

Participants who seek to present a poster during GL16 are invited to submit an English abstract between 300-350 words. The abstract should describe the project and/or related information products and services. The abstract should likewise include a title, name(s) of the presenter(s) and their full address information. Abstracts are an important source of information available prior to the conference that is accessible to both delegates and conference participants.

GL16 Conference Themes

- ☐ Public Awareness of Grey Literature
- ☐ Publishing and Licensing Grey Literature
- ☐ Open Access to Research Data
- ☐ Managing Change in Grey Literature
- ☐ Other Poster Topics/Themes:

Due Date and Format for Submission

Abstracts in MS Word should be emailed to conference@textrelease.com on or before **October 15th 2014**. The author will receive verification upon receipt. This will be accompanied by further guidelines for submission of PowerPoint slides, Conference Registration, etc. **GL16 will accommodate up to 20 physical posters (maximum size 30x40 inches or size A1), timely registration is a guarantee of your placement on the program!**

Sixteenth International Conference on Grey Literature



Grey Literature Lobby: Engines and Requesters for Change

The Library of Congress, Washington DC, USA

Federal Library Information Network, December 8-9, 2014

Conference Registration

- Think Tanks • Advocacy Groups • Research Teams • NGOs • IGOs • Foundations
- Coalitions • Lobbyists • Special Interest Groups • Task forces • Associations, Etc.

The Conference Registration Fee includes attendance during the Sessions, a copy of the GL16 Program Book, the full-text of the Conference Papers and PowerPoint Slides, as well as the conference pouch and participant badge. Also included are the lunches and refreshments during the breaks, and the post-conference tour.

Payment after October 1st 2014 add a \$75 surcharge

- ☐ \$595 Participant Fee ☐ \$495 Author Fee ☐ \$275 Day Pass ☐ \$100 Student Fee



I work for a **U.S. Government Agency** and am entitled to a 20% reduction on the fee ☐

I am a *GreyNet* Member and am entitled to a 20% reduction on the conference fee ☐

Registrant's Name: _____

Organization: _____

Address/P.O. Box: _____

Postal Code/City/Country: _____

Tel/Fax/Email: _____

Check one of the boxes below for your method of payment:

- ☐ Direct bank transfer to TextRelease, Account 3135.85.342 - Rabobank, Amsterdam, Netherlands
BIC: RABONL2U **IBAN:** NL70 RABO 0313 5853 42, with reference to GL16 and Registrant's Name

- ☐ MasterCard ☐ Visa Card ☐ American Express

Card No. _____ Expiration Date: _____

If the name on the credit card is not that of the registrant, print the name that appears on the card, below

Signature _____ CVC II code _____ (Last 3 digits on signature side of card)

Place _____ Date _____

Note: Credit Card transactions can be authorized by Phone, Fax, or Postal services. Email is not authorized.

Biagioni, Stefania graduated in Italian Language and Literature at the University of Pisa and specialized in data processing. She is currently a member of the research staff at the Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo" (ISTI), an institute of the Italian National Research Council (CNR) located in Pisa. She is head librarian and member of the ISTI Networked Multimedia Information Systems Laboratory (NMIS). She has been the responsible of ERCIM Technical Reference Digital Library (ETRD) and currently of the PUMA (Publication Management) & MetaPub, a service oriented and user focused infrastructure for institutional and thematic Open Access repositories looking at the DRIVER vision, <http://puma.isti.cnr.it>. She has coauthored a number of publications dealing with digital libraries. Her activities include integration of grey literature into library collections and web access to the library's digital resources, including electronic journals and databases. She is a member of GreyNet since 2005. Email: stefania.biagioni@isti.cnr.it

Giannini, Silvia graduated and specialized in library sciences. Since 1987 she has been working in Pisa at the Institute for the Science and Technologies of Information "A. Faedo" of the Italian National Council of Research (ISTI-CNR) as a librarian. She is a member of the ISTI Networked Multimedia Information Systems Laboratory (NMIS). She is responsible of the library automation software "Libero" in use at the CNR Research Area in Pisa and coordinates the bibliographic and managing activities of the ISTI library team. She cooperates in the design and development of the PUMA (Publication Management) & MetaPub, an infrastructure software for institutional and thematic Open Access repositories of published and grey literature produced by CNR. silvia.giannini@isti.cnr.it

Ilavská, Jana holds a master degree (Mgr.) from Comenius University in Bratislava, where she is undertaking a PhD. on Librarianship and Information Science. Her research interests are oriented on bibliometrics, scientometrics, infometrics and related topics including repositories. Since 2001 she has been working in Academic Library of Comenius University as system administrator and librarian. She is experienced in database systems, metadata and knowledge organization. She participates on several university or national projects, including building of Central Registry of Publication Activities of Slovak universities. Email: ilavska@rec.uniba.sk

Júlíusdóttir, Stefania is an Assistant Professor in the Department of Library and Information Sciences, Faculty of Human and Social Sciences at the University of Iceland. Her main fields of teaching and research are systematization of knowledge and information (i.e. cataloguing, classification and indexing), retrieval of knowledge and information, and the development of the subject field of collections of recorded knowledge and information (i.e. in libraries, records centres and archives). Moreover, her research field includes the effects of e-publishing on public access to publications (particularly grey literature), as stipulated in Icelandic legal deposit acts. The topic of her PhD thesis on which she is presently working, focuses on the effects of developments in the research areas named above on career opportunities and working conditions of staff in libraries, records centres and archives in Iceland. Her professional experiences include Director of the Library of the Icelandic Building Research Institute, Librarian in the National Department of the National Library of Iceland working on the National bibliography, Library Director (of public and school libraries) at the Ministry of Culture and Education, Library Director at the Directorate for Health in Iceland. Education 1974 BA-in Librarianship and Biology at the University of Iceland, MS-degree in Librarianship from Columbia University in the City of New York, PhD studies at Åbo Akademi in Finland and at the University of Iceland. Email: stj12@hi.is

Kopecká, Marcela studied Library and Information Science at Comenius University in Bratislava, Slovakia and she works as an information specialist in the Academic Library at the same university. She is a student of PhD degree at Comenius University in Bratislava. In her research she concentrates on e-publishing and the usage of institutional repositories in academic communication. She participates in research projects on information behaviour of PhD students. Other areas of interest include Open Access, repository development, metadata of open archives and open repositories, data interoperability and digital libraries. Email: kopecka@rec.uniba.sk

Lee, Seon-Hee is Senior Researcher at the Korea Institute of Science and Technology Information (KISTI). She has a master's degree in Library and Information Studies (MLIS) from the University of California, Los Angeles (1996) and a Master of Art degree in Philosophy from Ewha Women's University (1988). Her research interests include: collection development, grey literature, e-journals, information services, and collaborative digital reference services (CDRS). Email: wisdom@kisti.re.kr

Ondrišová, Miriam studied computer science at the Slovak University of Technology in Bratislava. She is a staff member of the Dept. of Library and Information Science of Faculty of Philosophy, Comenius University, Bratislava. As a teacher she reads lectures devoted to Bibliometrics, Database and Information Systems and Information technology. In research she works in a field of bibliometrics, information literacy and institutional repositories. miriam.ondrisova@fphil.uniba.sk

Pardelli, Gabriella was born in Pisa, graduated in Arts in 1980 at the Pisa University, submitting a thesis on the History of Science. Since 1984, researcher at the National Research Council, Institute of Computational Linguistics "Antonio Zampolli" ILC, in Pisa. Head of the Library of the ILC Institute since 1990, responsible for the Archives of the ILC Institute since 2005. Her interests and activity range from studies in grey literature and terminology, with particular regard to the Computational Linguistics and its related disciplines, to the creation of documentary resources for digital libraries in the Humanities. She has participated in many national and international projects including the recent projects:- BIBLOS: Historical, Philosophical and Philological Digital Library of the Italian National Research Council, (funded by CNR); - For digital edition of manuscripts of Ferdinand de Saussure (Research Programs of Relevant National Interest, PRIN - funded by the Ministry of Education, University and Research, MIUR). Email: gabriella.pardelli@ilc.cnr.it

Prost, Hélène is responsible for studies at the Institute of Scientific and Technical Information (INIST-CNRS). The different studies concern the evaluation of collections, document delivery, usage analysis, grey literature and open access to information. Expertise in statistical tools and knowledge in library information science allowed her to participate in various research projects and writing of several publications. Email: helene.prost@inist.fr

Savic, Dobrica is Head of the Nuclear Information Section (NIS) of the IAEA. He holds an MPhil degree in Library and Information Science from Loughborough University, UK, an MA in International Relations from the University of Belgrade, Serbia, as well as a Graduate Diploma in Public Administration, Concordia University, Montreal, Canada. Dobrica has extensive experience in the management and operations of web, library, information and knowledge management, as well as records management and archives services across various United Nations Agencies, including UNV, UNESCO, World Bank, ICAO, and the IAEA. His main interests are creativity, innovation and use of information technology in library and information services. He can be contacted at: d.savic@iaea.org

Schöpfel, Joachim is senior lecturer at the Department of Information and Library Sciences at the Charles de Gaulle University of Lille 3 and Researcher at the GERiCO laboratory. He is interested in scientific information, academic publishing, open access, GL and eScience. He is a member of GreyNet and euroCRIS. He is also the Director of the National Digitization Centre for PhD Theses (ANRT) in Lille, France. Email: joachim.schoepfel@univ-lille3.fr

Šušol, Jaroslav works as a professor in library and information studies, and the dean of the Faculty of Philosophy, Comenius University in Bratislava. He received his master degree in library and information science and English at Comenius University in 1986, license spéciale in information science and documentology at Université Libre de Bruxelles in 1992, and PhD in LIS with the thesis on the topic of retrospective conversion of library catalogues in 1993. He took part in several European research projects connected with network communication environment and its various aspects, like DELOS, CECUP, and CELIP. At present he is involved in research in the area of scholarly electronic communication, digital libraries, information and publishing behaviour and information ecology. Email: susol@fphil.uniba.sk

Notes for Contributors

Non-Exclusive Rights Agreement

- I/We (the Author/s) hereby provide TextRelease (the Publisher) non-exclusive rights in print, digital, and electronic formats of the manuscript. In so doing,
- I/We allow TextRelease to act on my/our behalf to publish and distribute said work in whole or part provided all republications bear notice of its initial publication.
- I/We hereby state that this manuscript, including any tables, diagrams, or photographs does not infringe existing copyright agreements; and, thus indemnifies TextRelease against any such breach.
- I/We confer these rights without monetary compensation and with the understanding that TextRelease acts on behalf of the author/s.

Submission Requirements

Manuscripts should not exceed 15 double-spaced typed pages. The size of the page can be either A-4 or 8½x11 inches. Allow 4cm or 1½ inch from the top of each page. Provide the title, author(s) and affiliation(s) followed by your abstract, suggested keywords, and a brief biographical note.

A printout or PDF of the full text of your manuscript should be forwarded to the office of TextRelease. A corresponding MS Word file should either accompany the printed copy or be sent as an attachment by email. Both text and graphics are required in black and white.

REFERENCE GUIDELINES

General

- All manuscripts should contain references
- Standardization should be maintained among the references provided
- The more complete and accurate a reference, the more guarantee of an article's content and subsequent review.

Specific

- Endnotes are preferred and should be numbered
- Hyperlinks need the accompanying name of resource and date; a simple URL is not acceptable
- If the citation is to a corporate author, the acronym takes precedence
- If the document type is known, it should be stated at the close of a citation.
- If a citation is revised and refers to an edited and/or abridged work, the original source should also be mentioned.

Examples

Youngen, G.W. (1998), Citation patterns to traditional and electronic preprints in the published literature. - In: College & Research Libraries, 59 (5) Sep 1998, pp. 448-456. - ISSN 0010-0870

Crowe, J., G. Hodge, and D. Redmond (2010), Grey Literature Repositories: Tools for NGOs involved in public health activities in developing countries. – In: Grey Literature in Library and Information Studies, Chapter 13, pp. 199-214. – ISBN 978-3-598-11793-0

DCMI, Dublin Core Metadata Initiative Home Page http://purl.oclc.org/metadata/dublin_core/

Review Process

The Journal Editor first reviews each manuscript submitted. If the content is suited for publication and the submission requirements and guidelines complete, then the manuscript is sent to one or more Associate Editors for further review and comment. If the manuscript was previously published and there is no copyright infringement, then the Journal Editor could direct the manuscript straight away to the Technical Editor.

Journal Publication and Article Deposit

Once the journal article has completed the review process, it is scheduled for publication in The Grey Journal. If the Author indicated on the signed Rights Agreement that a preprint of the article be made available in GreyNet's Archive, then browsing and document delivery are immediately provided. Otherwise, this functionality is only available after the article's formal publication in the journal.

An International Journal on Grey Literature

'WEIGHING UP PUBLIC ACCESS TO GREY LITERATURE'

AUTUMN 2014 – TGJ VOLUME 10, NUMBER 3

Using Google Search Appliance (GSA) to Search Digital Library Collections : A Case Study of the INIS Collection Search	123
<i>Dobrica Savić (Austria)</i>	
Grey Literature in European Commission Projects	133
<i>Sara Goggi, Gabriella Pardelli, Silvia Giannini and Stefania Biagioni (Italy)</i>	
Effect of Open API, NDSL Open Service (NOS) on Sharing Technical Reports in Korea	145
<i>Seon-Hee Lee and Mi Hwan Hyun (Korea)</i>	
Effects of legal changes and electronic publishing on the access to GL in Iceland	151
<i>Stefanía Júlíusdóttir (Iceland)</i>	
Scholarly publishing behaviour in Slovakia ... Are we ready for repositories?	163
<i>Jaroslav Šušol, Miriam Ondrišová, Jana Ilavská, and Marcela Kopecká (Slovakia)</i>	
Back to Grey: Disclosure and Concealment of Electronic Theses and Dissertations	173
<i>Joachim Schöpfel and Hélène Prost (France)</i>	

Colophon	118
Editor's Note	121
On the News Front	
Investigation into the use of Grey Literature in Evidence-Based Medicine: The role of the NHS Library and Information Services Professional	182
<i>Joanna Hooper, United Kingdom</i>	
GL16 Call for Posters 'Grey Literature Lobby: Engines and Requesters for Change'	184
GL16 Registration Sixteenth International Conference on Grey Literature, December 8-9, 2014, Library of Congress, Washington DC, USA	185
Advertisements	
EBSCO LISTA Full-Text	120
INIS, The International Nuclear Information System	122
Ila – Information International Associates, Inc.	172
About the Authors	186
Notes for Contributors	187