Fifteenth International Conference on Grey Literature

Slovak Centre of Scientific and Technical Information, Bratislava, 2-3 December 2013

THE GREY AUDIT: A FIELD ASSESSMENT IN GREY LITERATURE

Conference Proceedings

ISSN 1385-2316

Host and Sponsors

CENTRUM VEDECKO-TECHNICKÝCH INFORMÁCIÍ SR

LIBRARY OF CONGRESS FEDLINK

Nuclear Information Section
NIS - IAEA
International Atomic Energy Agency

NEW YORK ACADEMY OF MEDICINE 1847

cnrs Inist

JST

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Technical Library

SWINBURNE UNIVERSITY OF TECHNOLOGY

ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE "A. FAEDO"

KiSTi www.kisti.re.kr

GreyNet
Grey Literature Network Service

EBSCO PUBLISHING

GL15 Program and Conference Bureau
TextRelease
Javastraat 194-HS, 1095 CP Amsterdam, The Netherlands
www.textrelease.com ▪ conference@textrelease.com
Tel/Fax +31-20-331.2420

CIP

GL15 Conference Proceedings
Fifteenth International Conference on Grey Literature : The Grey Audit, A Field Assessment in Grey Literature, 2-3 December 2013 / compiled by D. Farace and J. Frantzen ; GreyNet International, Grey Literature Network Service. Amsterdam : TextRelease, March 2014. – 196 p. – Author Index. – (GL conference series, ISSN 1385 2316 ; No. 15)
CVTISR, EBSCO, FEDLINK-LC, INIST-CNRS, ISTI-CNR, KISTI, NIS-IAEA, NTK, and NYAM are corporate authors and associate members of GreyNet International. These conference proceedings contain the full text of 26 papers presented during the two days of plenary and poster sessions. The papers appear in the same order as in the conference program book. Included is an author index with the names of 78 contributing authors/coauthors along with their biographical notes. A list of more than 50 participating organizations as well as sponsored advertisements are likewise published in these conference proceedings.

ISBN: 978-90-77484-22-7

Foreword

THE GREY AUDIT

A Field Assessment in Grey Literature

The Fifteenth International Conference in the GL-Series provides the grey literature community an inclusive platform from which to assess developments in their field of information. Over the past two decades, since the very launch of this conference series, information science has been significantly impacted by social and technological developments. This gives sufficient cause for an audit in the field of grey literature – drawing upon accomplishments, assessing limitations, and projecting a sustained course of action.

A field assessment of grey literature extends well beyond library and information science, for it includes the assessment of grey literature produced and published in other sciences as well as government, business and industry. Information professionals and practitioners also become a part of this assessment, for it is they who carry out research in specific fields and make results available to their respective communities and wider public audiences.

The Grey Audit seeks to ascertain the validity and reliability of information and data produced in the grey circuit. It further seeks to measure the cost effectiveness of investing in grey literature both in material as well as human resources. The Grey Audit sets out to examine accepted standards applied in processing and distributing grey literature in an effort to identify guidelines for good practice that will be of benefit well into our 21st Century. Such examples of good practice will no doubt impact policy, which in turn will ensure future programs where grey literature is deployed. Twenty-six presentations from authors and researchers from 15 countries worldwide are harvested in this Proceedings.

Dr. Dominic J. Farace Amsterdam,
Grey Literature Network Service March 2014

GL15 Conference Sponsors

CVTISR, , Slovak Republic
Slovak Centre of Scientific and Technical Information

KISTI, Korea
Korea Institute of Science and Technology Information

Grey Literature Strategies, Australia
Swinburne University of Technology

EBSCO Publishing, USA

NIS-IAEA, Austria
Nuclear Information Section
International Atomic Energy Agency

ISTI, Italy
Institute of Information Science and Technologies
National Research Council, CNR

GL15 Conference Sponsors

FEDLINK, USA
Federal Library Information Network
Library of Congress

NTK, Czech Republic
National Library of Technology

NYAM, USA
New York Academy of Medicine

INIST-CNRS, France
Institut de l'Information Scientifique et Technique;
Centre National de Recherche Scientifique

JST, Japan
Japan Science and Technology Agency

GL15 Moderators and Conference Chair

Moderator Day One

Jens Vigen, Head Librarian CERN

For over a decade, Jens has been deeply involved in designing digital library services. In parallel to developing new services for members of the particle physics community, he has a particular interest for redesigning business models in the digital era for purchasing of library materials. Recently his activities have been strongly focused on establishing models for open access journal publishing. Before joining CERN, Jens held a position at the library of the Norwegian University of Science and Technology. In addition to his library qualifications he has a master degree in civil engineering; geodesy and photogrammetry. jens.vigen@cern.ch

Conference Chairman

Ľubomír Bilský, Head Officer CVTISR

Ľubomír started his professional career in 2001 at the Business and Innovation Centre Bratislava, where he was responsible for international projects implementation. He was an active member of the Innovation Relay Centres (IRC) Network dealing with the support of transnational research and technological co-operation development. In 2008, he became the head of Innovation section within the Enterprise Europe Network Slovakia consortium, an EC initiative supporting business, innovation and research co-operation of research organisations and SMEs. At Slovak Centre of Scientific and Technological Information (CVTI SR), he has been responsible for preparation and implementation of several national and international projects focused on support of scientific community in Slovakia, including active promotion of science and technology in society. He also co-ordinates activities related to building the national infrastructure for technology transfer support in Slovakia.
lubomir.bilsky@cvtisr.sk

Moderator Day Two

Dobrica Savić, Head NIS-IAEA

As head of the Nuclear Information Section, Dobrica is responsible for the International Nuclear Information System (INIS), the IAEA Library, and the IT support group. He holds a MPhil 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 long-term sustainability of information services, democratization of scientific and technical information, and the practical application of information technologies.
d.savic@iaea.org

Table of Contents

GL15 Program Committee

	Ján Turňa **Chair** Marta Dušková **Co-Chair** Ľudmila Hrčková **Co-Chair** Slovak Centre of Scientific and Technical Information, CVTI SR	Slovak Republic
	Danielle Aloia New York Academy of Medicine, NYAM	United States
	Stefania Biagioni Institute of Information Science and Technologies, ISTI-CNR	Italy
	Kiseok Choi Korea Institute of Science and Technology Information, KISTI	Korea
	Daniela Luzi Institute of Research on Population and Social Policies, IRPPS	Italy
	Petra Pejsova National Library of Technology, NTK	Czech Republic
	Dobrica Savic Nuclear Information Section, NIS-IAEA	Austria
	Joachim Schöpfel University of Lille 3	France
	Roberta Shaffer Library of Congress, LC	United States
	Christiane Stock Institut de l'Information Scientifique et Technique, INIST-CNRS	France
	Dominic Farace Grey Literature Network Service, GreyNet International	Netherlands

OPENING SESSION

SESSION ONE – TECHNOLOGY ASSESSMENT

SESSION TWO: SUSTAINING GOOD PRACTICES

SESSION THREE – RESEARCH AND DATA

SESSION FOUR – TOWARDS INFORMED POLICIES

POSTER SESSION

Slovak Centre of Scientific and Technical Information **SCSTI**

Achieve your goals with us

CENTRUM VEDECKO-TECHNICKÝCH INFORMÁCIÍ SR

INFORMATION SUPPORT OF SLOVAK SCIENCE

SCIENTIFIC LIBRARY AND INFORMATION SERVICES

- technology and selected areas of natural and economic sciences
- electronic information sources and remote access
- depository library of OECD, EBRD and WIPO

SUPPORT IN MANAGEMENT AND EVALUATION OF SCIENCE

- Central Registry of Publication Activities
- Central Registry of Art Works and Performance
- Central Registry of Theses and Dissertations and Antiplagiarism system
- Central information portal for research, development and innovation - CIP RDI >>>
- Slovak Current Research Information System

SUPPORT OF TECHNOLOGY TRANSFER

- Technology Transfer Centre at SCSTI
- PATLIB centre

POPULARISATION OF SCIENCE AND TECHNOLOGY

- National Centre for Popularisation of Science and Technology in Society

IMPLEMENTATION OF PROJECTS

- National Information System Promoting Research and Development in Slovakia – Access to electronic information resources - NISPEZ
- Infrastructure for Research and Development - the Data Centre for Research and Development - DC VaV
- National Infrastructure for Supporting Technology Transfer in Slovakia - NITT SK
- Fostering Continuous Research and Technology Application - FORT
- Boosting innovation through capacity building and networking of science centres in the SEE region - SEE Science

www.cvtisr.sk
Lamačská cesta 8/A, Bratislava

Slovak Centre of Scientific and Technical Information: Who we are and what we do

Marta Dušková
Slovak Centre of Scientific and Technical Information (CVTI SR), Slovak Republic

Abstract
In this year Slovak Centre of Scientific and Technical Information (CVTI SR) celebrates the 75th anniversary of its establishment. In course of its existence our institution went through various organizational changes and we can say that today it is not only a library but also an information centre supporting science in Slovakia, participating in many national and international projects.
The main objective of paper is to inform the international grey literature community about the mission, objectives and activities of CVTI SR as an institution, which is the only one in Slovakia comprehensively dealing with grey literature.

Four Ways – Four Activities
Current activities of CVTI SR are implemented in four main areas: scientific library, support of science, popularisation of science, and projects.

First way: Scientific Library
CVTI SR is a specialized public scientific library that helps to develop science, technology and education by creating and administrating library and information funds and by providing librarian and information services to wide expert public. The library acquires, stores and presents all types of national and foreign documents, including standards, norms, patents, corporate literature, grey literature and also documents of depository libraries of OECD, EBOR, WIPO and the EU in printed and electronic forms and provides for comprehensive library, information and search services for science and research in Slovakia and also for public.
Our institution as a specialized public scientific library is under *Library Act no. 183/2000 Coll.*, a depository library and a coordinator for processing documents normally issued in smaller amounts of research institutes, universities, international organizations, public bodies and other legal and individual entities, which aren't issued by publishers in traditional distribution network. It follows, that it is responsible for the collection, processing and storage of grey literature in the Slovak Republic.
In this way, how CVTI SR is involved in grey literature issues, is the collection, storage and disclosure of grey literature, which our institution acquires by the Act no. 535/2003 Coll. for mandatory copies of periodicals, non-periodical publications and audio-visual works. To most recent activity (from January 2013) belongs creating a depository library of grey literature for Association of Slovak Scientific and Technological Societies directly in CVTI SR.

Second way: Support of Science
Support of science presents the second part of activities in which CVTI SR participates today. Information support to science is secured by operating several national systems, each of them also includes elements of grey literature:

Central Information Portal for Research, Development and Innovation (CIP VVI) offers information connected with Government support of science and financing, relevant official documents, defined statistical data, implemented EU regulations, international scientific cooperation, results of research and development, research projects tenders.
CIP VVI contains three categories of information: Websites devoted to popularisation of science and technology in the society; Websites devoted to support of science and research in Slovakia and in the EU, possibilities how to finance them and presentation of their results; and information system database on science, research and innovations.
CIP VVI was in 2010 granted the *Microsoft Industry Awards* in category Best Solution for State Administration, Self-administration and Academic Sphere and in 2009 we acquired the *Inforum* award as the best product / service connected with electronic information resources.

Slovak Current Research Information System (SK CRIS) was introduced in operation in 2013 as a new information system for the area of science and research. SK CRIS contains a database of information on projects financed by public resources, their results, and it also contains a register of research and development organizations and a personal database of researchers. The database was drafted in accordance with the European standards (CERIF data format), that according to requirement of the EU Commission is managed by the international organization *euroCRIS*.

Central Registry of Publication Activity (CREPČ) and ***Central Registry of Art Work and Performance (CREUČ)*** were established as a development project of the Ministry of Education, Science, Research and Sport of the Slovak Republic in years 2007 – 2008. Constantine Philosopher University in Nitra, University of Žilina and the company SVOP, spol. s r.o. were researchers of the project. All public and national universities and also defined private universities in the Slovak Republic are also contributors to CREPČ and CREUČ.

The objective of the registers is the complex automatised evidence of publication and art activities; creating a unique information resource for experts and lay public; providing data for calculating Government subsidies to the universities in the Slovak Republic.

These systems are supported by legislation: *Act 131/2002 on Universities as amended; Ordinance of the Ministry of Education, Science, Research and Sport No. 456/2012 on Central Register of evidence of publication activities and Central Register of evidence of art activities and the Directive of the Ministry of Education, Science, Research and Sport No. 13/2008-R on Bibliography Registration and Categorization of publication activities, art activities and responses.*

Central Registry of Theses and Dissertations (CRZP) a Antiplagiarism System (APS) are two closely connected cooperating systems. CRZP serves as the central repository for long term storage of bachelor, diploma, doctor's, dissertation and habilitation theses and the APS secures their originality check. Technically are the CRZP a APS systems construed as the systems of co-working servers: application server, storage server, database server and antiplag server.

Currently there are 33 universities connected to the CRZP/APS system. Approximately 75 000 final and qualification theses enter the CRZP per year, and currently there are approximately 300 000 theses stored. More than 5 million documents were downloaded at the Internet for comparison and plaque detection in volume of approximately 2 TB data. Comparing of one document with the whole current corpus including processing metadata and generating the PDF protocol lasts 10 seconds in average.

The CRZP/APS project was awarded in the Best Project competition in area of Informatization of Public Administration at the ITAPA 2011 conference, where in the category of new services CVTI SR acquired 2nd price. In 2013 CVTI SR succeeded in the European wide competition *"European Prize for Innovation in Public Administration"* called by the EU Commission with the CRZP/APS Project and was awarded 100 000 EUR for the 1st prize in category *"Initiatives in area of Education and Research"*.

Third way: Popularisation of Science

In 2007 the ***National Centre for Popularisation of Science and Technology (NCP VaT)*** was established at the CVTI SR. Its task is to secure activities connected with popularization of science and technology in society.

Foundation of NCP VaT created an administrative – organization background for the whole academic area (universities, Slovak Academy of Sciences, state research and development organizations, research workplaces in industries, non-profit research and development organizations) to present results in research and development, problems in organizing research and development with maintaining and acquiring human resources in research and development, modernization and creating infrastructure; for young people, who are interested in research and development, media representatives, entrepreneurs, expert and wide public.

NCP VaT within their activities organizes various types of events: a series of events "*Science in Center"* (series of events for seniors, for students, with renowned Slovak scientists concerning interesting themes from the world of scientific – technical discoveries), *Week of Science and Technology in Slovakia, Researchers Night, Science Spectrum*) series of TV documents on work and success of Slovak researchers, their extraordinary disclosures), *Laboratory* (cyclic discussion program on science in Devín Radio). NCP VaT issues an electronic online magazine *Scientific Kaleidoscope* and operates the *Central Information Portal and Information System for Research, Development and Innovations, SciTech Navigator* (catalogue of free Internet resources for all areas in science and technology) and *SciTech blog* (non-formal online space for discussions on science and technology).

Fourth way: Projects

In recent years the activities of CVTI SR concentrated also to implementation of national and international projects. The following projects belong to the biggest and the most significant ones:

National Information System for Supporting Research and Development in Slovakia – Access to Electronic Information Resources (NISPEZ)

The strategic goal of the national project is to implement an effective system for information support for research and development in Slovakia. This information system is based on electronic information resources and its effectiveness is achieved by tools for the management and administration of electronic resources. Created within the framework of the project, are a database of Slovak electronic resources for R&D and SK CRIS, a new R&D information system in compliance with European standards.

Infrastructure for Research and Development – Data Centre for Research and Development (DC VaV)

Strategic goal of project is to establish the Data Centre for Research and Development with sufficient capacity to store and process the complex information essential for R&D in Slovakia and to provide auxiliary services. An infrastructure for electronic communications on R&D are also included in these services. An infrastructure for instant access to services and safety of the operation is of prime concern.
The Data Centre for Research and Development at Žilina University in Žilina with its back-up centre in Bratislava was officially opened in June 2010. It is a data storage facility equipped with high-speed computing technology which is capable of processing particularly large-capacity data sets.
The Department of Digitizing was established at the CVTI SR within the framework of the project in May 2010. It is equipped with modern digitizing technology with subsequent processing of the digitized documents. It has the capacity to effectively digitise various types of documents including large-scale maps and plans to meet the requirements of the scientific community.

National Infrastructure for Supporting Technology Transfer in Slovakia (NITT SK)

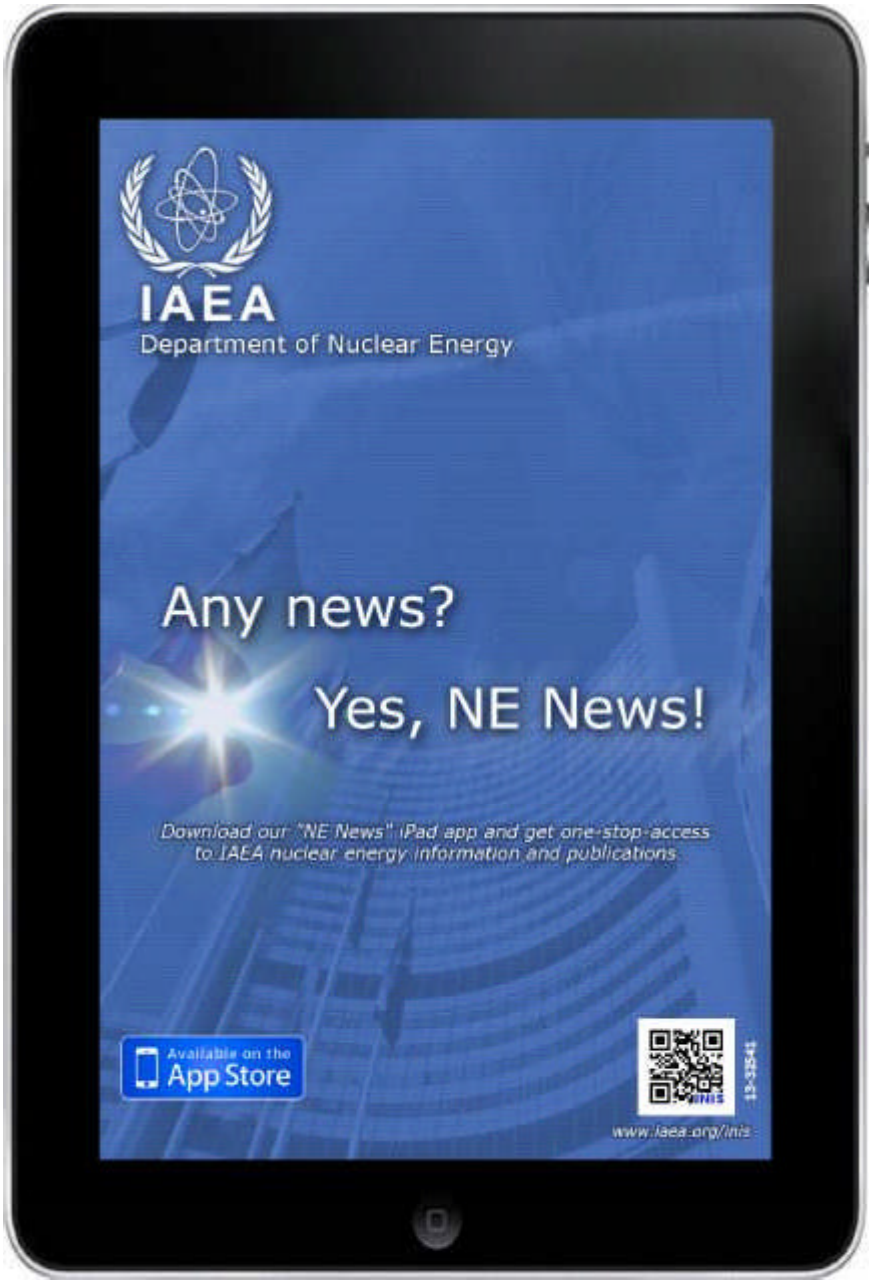
The creation and implementation of a national infrastructure to support the technology transfer resulting from research and development in economic and social practices are the main objectives of project. The national system will support those R&D activities that respond to the needs of the business sector. As a consequence, an increased number of achievements in R&D and innovation and technologies will be put into practice. The project will facilitate the creation of long-term partnerships between the academic and business sectors and will contribute to the sustainable development of the whole of society. The Technology Transfer Centre will be established at the CVTI SR to ensure that the support for technology transfer is systematic at the national level. As part of the National System for Supporting Technology Transfer (NSPTT), the National Portal for Technology Transfer - NPTT is operated within the framework of the national project implementation. The NPTT is a website where all the important ongoing information in technology transfer in Slovakia can be found and it is also the site from which expert support services, relating to the process of technology transfer provided to academic and research organisations, can be accessed.

Promotion of Science and Technology in Slovakia (PopVaT)

The aim of the project is to improve the perception and status of science in society through the promotion of science and technology towards the wider public, including the youth.
Specific objective of the project is the enhancement of public awareness, including the youth, about the importance of science and technology, as well as the enhancement of awareness of the scientific community about the importance of science promotion.
The main expected outcome of the project is the change in the perception and status of Slovak science and technology in society, through the promotion of science and technology towards the three main target groups: the youth, the general public and the scientific community. Implementation of the project shall significantly contribute to the increase of public awareness related to scientific activities, research results obtained and their possible application into the practice.

References

Centrálny informačný portál pre výskum, vývoj a inovácie. CVTI SR. [online]. [cit. 2013-11-03]. Avilable at:: https://www.vedatechnika.sk/SK/Stranky/default2.aspx
Centrálny register záverečných prác. CVTI SR. [online]. [cit. 2013-11-10]. Avilable at: http://www.crzp.sk/
Centrum vedecko-technických informácií SR. CVTI SR. [online]. [cit. 2013-11-13]. Avilable at: www.cvtisr.sk
Národná infraštruktúra pre podporu transferu technológií na Slovensku – NITT SK. CVTI SR. [online]. [cit. 2013-11-05]. Avilable at: http://nitt.cvtisr.sk/
Národné centrum pre popularizáciu vedy a techniky v spoločnosti. CVTI SR. [online]. [cit. 2013-11-05]. Avilable at: http://ncpvat.cvtisr.sk/
Národný informačný systém podpory výskumu a vývoja na Slovensku – prístup k elektronickým informačným zdrojom. CVTI SR. [online]. [cit. 2013-11-05]. Avilable at: http://nispez.cvtisr.sk/
Infraštruktúra pre výskum a vývoj: Dátové centrum pre výskum a vývoj. CVTI SR. [online]. [cit. 2013-11-15]. Avilable at: http://dc.cvtisr.sk/
Portál CREPČ/CREUČ. CVTI SR. [online]. [cit. 2013-11-05]. Dostupné z: http://cms.crepc.sk/
SK CRIS: Informačný systém o vede a výskume. CVTI SR. [online]. [cit. 2013-11-05]. Avilable at: https://www.skcris.sk/portal/


IAEA
Department of Nuclear Energy
Any news?
Yes, NE News!
Download our "NE News" iPad app and get one-stop-access to IAEA nuclear energy information and publications
Available on the App Store
www.iaea.org/inis

Digital Preservation at International Nuclear Information System (INIS)

Dobrica Savic and Germain St-Pierre
Nuclear Information Section, IAEA, Austria

Abstract
Since its creation in 1970 until 1996 (INIS, 2010), the International Nuclear Information System (INIS)[i] collected and converted to microfiche over 312 000 non-conventional literature (NCL) reports received from member states and international organizations. The microfiche collection contains over 1 million items, with an estimated total of 25 million pages of full-texts.
In 1997, the INIS Secretariat replaced the microfiche-based production system with an imaging system to process and to disseminate all NCL documents in electronic format. That marked the beginning of digital preservation efforts that still continue today.
This paper provides an overview of the digital preservation practices and the technical infrastructure of INIS. It describes the technical processes, the standards in place, the hardware and software used, as well as all practices related to scanning, quality control, OCR, preservation and storage.

Technical Infrastructure
The INIS digital preservation technical infrastructure has evolved on a regular basis since the beginning of digital imaging activities. Initial period known as INIS Imaging System lasted from 1997 to 2003, to be followed by a new INIS Imaging System (INISIS2K) period which lasted from 2003 to 2009. Current technical infrastructure was introduced in 2010.

INIS Imaging System (INISIS) - 1997 to 2003
In 1997, Jouve Systems was selected as a full-scale imaging system to process and disseminate INIS NCL in electronic format (INIS, 1999). This "cradle-to-grave" image-based solution replaced the microfiche-based production system which had been in place at the INIS Secretariat since 1970. The following modules were already part of the original design: workflow monitoring, black and white scanning, image import, image enhancement, quality control, link creation using barcode recognition, link validation against INIS bibliographic metadata and INIS rules, cumulative index creation as well as CD-ROM production according to the INIS NCL Viewer specifications (INISIR). Originally, only the TIFF Group 4 format was supported. In 2002, support for incoming documents in PDF was added, although the Jouve system was phased out only in 2003.

INIS Imaging System (INISIS2K) - 2003 to 2009
A study carried out in 2000 by Doculabs recommended building a new INIS Imaging System (INISIS2K) on one of the following "off-the-shelf" 32-bit information capture systems: Kofax Ascent Capture or ActionPoint InputAccel (now part of the EMC-Captiva family))[ii]. InputAccel (IA) was selected, mainly because of its powerful open architecture technology that allowed customization and system integration with Open-Text Livelink[iii], the IAEA standard Document Management System. InputAccel also met new requirements such as colour scanning, optical character recognition (OCR) and output to PDF.
Replacement of the INISIS imaging system led to a significant improvement in the production cycle, which was synchronized with the bibliographic database production. All documents were output in PDF and those in Western European, Cyrillic and Slavic scripts were OCRed (INIS, 2004).
From the beginning, INISIS2K was conceived and implemented as one of the components of a larger system, a completely overhauled INIS Processing System (IDPS) based on Livelink technology. All tasks, from the initial imaging request sent to the InputAccel server until the ingestion of its PDF output into the document repository, were monitored through Livelink. This was also the case for the quality control of bibliographic data, the ingestion of NCL input submitted by the National Centres in PDF format, the migration of all new records to the INIS Online Database, and finally for the preparation of an ISO image for distribution of the full-texts on CD-ROM.
In 2006, in order to streamline workflow, improve efficiency and free resources for other activities, the INIS Secretariat issued revised 'Guidelines on How to Submit Full-Text of Non-Conventional Literature (NCL) to INIS' (INIS, 2006). The INIS National Centres were strongly encouraged to submit their NCL input directly in PDF and the response from Member States was favourable.
Three new priorities were identified: the digitization of the INIS microfiche collection, the conversion to PDF of all the documents scanned and distributed in TIFF between 1997 and 2003, and the online access to full-texts via the INIS Online Database.

Although highly efficient when introduced in 2003, InputAccel lacked flexibility when it came to the development of workflows tailored for other digitization projects. The maintenance of this modular client/server application was also very expensive and required significant effort from the Systems Development and Support Group (SDSG). Finally, incompatibility of the communication module with Livelink was found during testing of InputAccel v.5.3. This made the migration to this new platform impossible without additional expensive developments. INIS decided to stop the maintenance contract for InputAccel at the end of 2009 and abandoned the system with the migration of all desktops to Windows 7 in 2010.
During this period, the INIS imaging infrastructure consisted of 6 scanning workstations, 3 servers, 4 high performance scanners, 2 flatbed scanners, 1 high performance microfiche scanner and 1 digital camera. The technical characteristics are indicated in the table below.

Table 1: Imaging Infrastructure 2003 – 2009

Scanner	Type	Paper size	Resolution (dpi)	Bit-in-depth	Speed (A4, 200 dpi)	ADF Page capacity
Fujitsu fi-5750c with VRS Pro	Colour; ADF/flatbed	A8 – A3 Up to 34 inches	50 to 600	24	110 p/min (simplex) 55 p/min (duplex)	200 p
Fujitsu M4099D	B&W; ADF	A7 – A3	200, 240, 300, 400	10	90 p/min (simplex) 180 p/min (duplex)	1000 p
Fujitsu M3099GX	B&W; ADF	A7 – A3	200, 240, 300, 400	8	60 p/min (simplex) 120 p/min (duplex)	1000 p
Fujitsu M3099G	B&W; ADF	A5 – A3	200, 240, 300, 400	8	55 p/min (simplex) 110 p/min (duplex)	500 p
Kodak i260	Colour; ADF/flatbed	A5 – A3	Up to 600 Optical Resolution 300	16-48	50 p/min (simplex) 100 p/min (duplex)	150 p
SunRise 2000	Microfiche scanner	A0-A4 reductions 7x-50x	CCD 3600-8800 True Resolution		Up to 2500 frames/hr	

Current Technical Infrastructure

A complete re-evaluation of the technical infrastructure was carried out in 2010, in line with the implementation plan of the 'Desktop 2010' project developed by the IAEA Division of Information Technology (MTIT) (INIS, 2011). An important goal of this project was to ensure security and supportability of all computer systems of the Agency network.
Windows 7 compliance of all equipment and software applications had to be verified through testing prior to the deployment of this new platform. Also, an important reduction in space requirement was an expected outcome of this exercise.
The 3 Fujitsu black and white SCSI scanners, the Kodak i260, the InputAccel system and some small utilities failed this compliance test. Also, several old workstations did not meet the minimum requirements and had to be replaced.
New computers with fast quad-core processors supporting multithreading and multitasking were procured. The number of scanners was reduced to two, both of them supporting color, greyscale and black and white scanning.

Software

The following software and applications are currently used for digitization at INIS:
Techsoft PixEdit v.7.11.18: PixEdit was introduced in the imaging workflow in 2000. It is primarily used for its advanced image editing capabilities. This flexible application gradually proved to be an excellent scanning utility. Since the discontinuation of the InputAccel system in 2010, PixEdit is the main scanning application. Five seat licenses are currently available.
ABBYY FineReader 11 Corporate Edition: FineReader is used for Optical Character Recognition (OCR). It can process mono or multilingual documents, supports different alphabets including Cyrillic languages and offers an accuracy level of close to 98%. ABBYY policy for this product is to release a new version each year. Version 11 was bought in 2011 together with an upgrade assurance to Version 12 in 2012.
Adobe Acrobat X Professional is used for OCR of Chinese (Simplified), Japanese and Korean, as well as for document optimization and conversion to PDF/A[iv], when applicable.
Kofax Virtual ReScan (VRS) + Kodak Perfect Page: Both technologies have hardware and software components that reduce the need for post-scanning image enhancement.

Hardware

Scanners - One of the most important elements in a digitization project is the selection of the appropriate image capture devices, as scanners have great impact on image quality. The choice of

equipment depends on a number of factors, including the format, size and condition of the material that will be digitized.
Several types of digitization equipment exists, i.e. flatbed scanners, sheet-fed scanners with automatic document feeder (ADF), drum scanners, open book scanners, digital cameras, and film scanners.
INIS quality scanners are calibrated and maintained regularly. Special methods, including Scanner Test Charts, are used to check image resolution, dynamic range mapping, as well as photographic tone and color reproduction.
There are currently 2 colour scanners with automatic document feeder (ADF) and flatbed, as well as 2 high performance microfiche scanners. The technical characteristics are indicated in the table below.

Table 2:- INIS Scanner Specifications

Scanner	Type	Paper size	Resolution (dpi)	Bit-in-depth	Speed (A4, 200 dpi)	ADF Page capacity
Fujitsu fi-5750c with VRS Pro	Colour; ADF/flatbed	A8 – A3 Up to 34 inches	50 to 600	24	110 p/min (simplex) 55 p/min (duplex)	200 p
Kodak i1440	Colour; ADF/flatbed	A5 – A3	Up to 600 Optical Resolution 300	16 - 48	50 p/min (simplex) 100 p/min (duplex)	150 p
SunRise 2000	Microfiche scanner	A0-A4 reductions 7x-50x	CCD 3600-8800 True Resolution		Up to 2500 frames/hr	
SunRise Apollo	Microfiche scanner	A0-A4 reductions 7x-50x			Up to 3600 frames/hr	

Computers - Careful consideration was given to the following points when selecting PCs dedicated to digitization work: fast central processor unit (CPU), sufficient random access memory (RAM), fast data transfer rate between components, large disk storage capacity, suitable interface, as well as high quality audio and video cards.
Monitors - Large display monitors provide better viewing and image evaluation. As each type, size and quality of monitor interprets and displays values differently, special care is devoted to their adjustment and calibration. Four quality control workstations and the 2 scanning workstations are equipped with widescreen monitors (30-inch LCD monitors, model LP 3065 from Hewlett-Packard).

General digitization principles at INIS
INIS aims to ensure a consistent, high level of image quality, interoperability and accessibility of digitized materials, as well as long-term preservation for future generations in Member States. To achieve these goals, INIS developed some general principles based on Cornell University's digital imaging tutorial (http://www.library.cornell.edu/preservation/tutorial/index.html), and adjusted them to INIS requirements. The current workflow is based on these principles and described in detail later. The principles can be summarized as follows:

1. benchmarking for image quality and resolution
2. scanning at the level appropriate to the content of the original source
3. digitization of 1st generation material, if available, in order to achieve best possible image quality
4. creation and storage of a master image file
5. use of format and compression techniques that conform to standards (avoiding proprietary formats)
6. creation of backup copies
7. storage of digital files in an appropriate environment
8. off-site storage of the collection
9. metadata for digital resources
10. integration of image files with bibliographic metadata in the INIS Collection

Image Creation Process
The process of initial capture or conversion of a paper or microfiche based document or object into digital form is known as image creation. Based on the experience gained over the years and through benchmarking, INIS dedicates special attention to the physical nature of the documents to be digitized and applies different measures. Collections available at Member States differ in the ways they are created, used and accessed. The quality and condition of the original material will have a direct impact on the digitization approaches. Therefore, INIS applies the principle to scan at a level that matches the information content of the original.

Before starting a digitization project, it is crucial to obtain copyright permissions from copyright holders. The Intellectual Property Rights laws are comprehensive and complex, and the progress in today's online environment presents serious challenges for copyright compliance.
In this respect, INIS relies on INIS Member States to ensure that appropriate permission is obtained before the full-text of a publication is sent to INIS for inclusion in the Collection.
It is essential that each stage of the digitization process is planned ahead and an appropriate workflow is established. It is not a simple task to create an effective and efficient ***digitization workflow***. However, if properly planned it will support staff performance and enable high quality work. The stages of workflow at INIS are the following:

- Benchmarking
- Source material types
- Preparation
- Scanning
- Quality control
- Image enhancement
- File formats
- Compression
- File naming convention
- Optical character recognition (OCR)
- Storage
- Preservation planning
- Metadata creation

Benchmarking
INIS considers benchmarking for digital capture the first and most important step of the digitizing effort. The results of benchmarking considerably affect all further steps (scanning, enhancement, format, etc.). The purpose of benchmarking is to define and clarify the following:

- Can the informational content of the original material be adequately captured in digital form?
- Does the physical format and condition of material correspond to digitizing requirements?
- What is the type of material to be digitized?
- Which resolution should be applied?
- At which bit-depth?
- Which compression parameters should be set?
- What is the estimated accuracy level for OCR?
- Other considerations?

Source material types
The variety of source material may be categorized, but not limited to:

- Printed text/simple line art
- Rare or damaged printed text
- Manuscripts
- Maps, architectural drawings
- Halftones
- Continuous tone
- Microformats
- Mixed

The majority of the material digitized at INIS is text-based containing illustrations, graphics, photos (black & white, colour), as well as oversized materials with fine details, line drawings, etc., falling mainly into the above cited *Printed Text* and *Mixed* categories. The category of *Printed Text* can be described as distinct edge-based representation that is cleanly produced, with no tonal variation, such as a book containing text and simple line graphics. Documents containing two or more of the categories listed above, such as illustrated books, can be defined as *Mixed.*

Preparation
Good document preparation facilitates scanning and ensures quality results. Materials to be scanned need to be prepared in the following manner:

- Physically (unbinding, removing of staples and clips, separation when glued, etc.);
- Structurally (adding/removing barcodes, separating chapters, sections, parts, covers, etc.);
- According to specific characteristics, e.g. size, thickness, quality (glossy/mat), condition of paper, etc.

Inadequate document preparation can result in paper jams inside the scanner and lead to irreparable damage to original documents. In order to unbind documents in an efficient and safe way, INIS utilizes a professional cutting machine from IDEAL (Model 4850-95).

Scanning

Capture modes: It is important to keep in mind that different capture methods are needed depending on the physical form of the original. Capturing is mostly performed in these 3 modes:

- Bitonal (1 bit per pixel) – represents two tones: black and white; best suited to high contrast documents such as printed text.
- Greyscale (8 bits per pixel) – represents 256 shades of grey; best suited to continuous tone documents such as black and white photographs. However, older photos (e.g. sepia tones) may provide better results when captured in colour.
- Colour (24 bits per pixel) – represents 16 million colours & shades of grey; suited to documents with continuous tone colour information.

'*Pixel*' stands for picture elements which make up an image. Each pixel can represent a number of different shades or colours depending on the storage space allocated to it.

Optical Resolution: The optical resolution determines the quality of an image. It is normally expressed in scanner specifications as 'dots per inch' (DPI) or 'pixels per inch' (PPI) and refers to the number of pixels (dots) captured in a given inch. Increasing the resolution enables capturing of finer detail. However, it results in a larger file size. To determine the resolution necessary to capture all significant details present in the source document, Cornell University developed a formula called 'Digital Quality Index' (QI). This formula can be used as guidance for calculating the optimal scanning resolution**.** More information is available on Cornell's Web site at: http://www.library.cornell.edu/preservation/tutorial/

Bit depth: The amount of information that a sensor in an array can capture is represented by the 'bit depth'. Greater bit depths result in a more accurate digital representation of the original. The final decision about resolution and bit depth depends on the goal of digitizing**.**

INIS applies a resolution range of 300 – 600 dpi for bitonal scanning to documents of A4-A5 size, and 200 – 300 dpi with 8 bit depth (256 colours/tones) for greyscale/colour scanning.

Quality Control (QC)

Quality Control is as an integral part of the digitization process in order to retain value, utility and integrity of the resources. QC consists of a set of procedures and techniques to verify the quality, accuracy and consistency of digitized material. QC is conducted by visual inspection of images on-screen with concentration on the resolution, colour, tone, and appearance. It is important to mention that this assessment may be highly changeable depending on the viewing environment and the characteristics of the monitors. INIS applies a wide range of QC measures to ensure that quality expectations are met. During the QC process, INIS verifies accuracy and completeness of components, data integrity, metadata correctness, form and validity, as well as correct matching of metadata and image files. In this context, the 'checksum' algorithm serves to ensure the authenticity and integrity of digitized files. It is essential to verify that the number and order of bytes in a file remain the same after moving, copying, transferring, burning or other actions. In addition to the checksum, INIS also compares the number of pages of the original with the digitized product to ensure the completeness of digitized documents.

Image enhancement

Image enhancement is any process that is applied to the raw scan to improve quality or legibility of the resource. INIS applies several procedures and techniques to verify the quality, accuracy, consistency and integrity of digital products, including despeckling, deskewing, noise reduction, black border removal; colour and tone adjustment, etc.

File formats

There are several standard file formats which vary in terms of resolution, bit-depth, colour capabilities, etc. Although there is no clearly recommended archival format in use today, preference must be given to 'non-proprietary' formats. INIS stores master digital images in **TIFF** Group IV which offers longevity and production of a range of delivery versions (e.g. for screen, for print, for web access). For purposes of delivery to Member States, electronic exchange with customers, users, and access via the INIS Online Database, files are converted to PDF (Portable Document Format) and compressed. **PDF** is one of the most frequently used file formats to preserve electronic documents and ensure their survival for the future. Recently, the International Organization for Standardization (ISO) released the full PDF specification as 'ISO 32000-1:2008'. **PDF/A:** The PDF/Archival (PDF/A) standard aims to enable the creation of PDF documents whose visual appearance will remain the same over the course of time. This standard was adopted by the International Organization for Standardization (ISO) in autumn 2005 and published as 'ISO 19005-1:2005'. INIS is considering adopting this standard to achieve preservation and long-term archiving of the Agency's and Member States' nuclear information resources. (that last sentence should perhaps be written a bit differently – not ideal the way it is now written…)

Compression

Compression algorithms are used to reduce image file size for storage, processing and transmission. There are two compression techniques, i.e. 'lossless' and 'lossy'[v]. When lossless compression is applied, the space needed for the storage of an image file is reduced without loss of data. During lossy compression, the least significant information is averaged or discarded. Uncompressed files or compressed files using the lossless compression technique are clearly preferred. There are several standards, as well as proprietary compression software available to create images for web delivery. INIS has chosen the JBIG2 standard for web optimization of black/white resources, and JPEG for colour digital resources.

JPEG2000 uses wavelet compression to achieve small but high quality images and is increasingly being used as repository and archival image format. INIS is considering JPEG2000 as a possible alternative for image delivery.

File naming conventions

For system compatibility and interoperability, it is important to follow an established file naming convention. Unique file names assure consistency and easy retrieval of resources.

Optical Character Recognition (OCR)

In order for an image of a printed text to become searchable as electronic text, raster images are processed with an OCR program to be translated to machine editable text.

For INIS digitization projects, the creation of 'searchable full text' has been defined as the primary objective. The quality and condition of the original material will have a direct impact on the OCR result.

INIS uses ABBYY FineReader, an Optical Character Recognition (OCR) software that allows users to convert paper documents, PDF files, and various images including photographs taken by a digital camera to editable formats for changing and repurposing. Close to 98% accuracy is reached at character level when applying OCR to raster images of text printed in Latin and Cyrillic characters.

Recent tests have provided satisfactory results using Adobe Acrobat Professional 8.0 for OCR of documents in Chinese (Simplified), Japanese and Korean. Tests with ABBYY FineReader Pro9 for Hebrew and Thai also provided good results. Further tests are being performed to identify suitable tools for the Arabic language.

Latest developments in OCR technology include recognition of document structure known as 'Logical Form Recognition' (Omnipage 16) or 'Adaptive Document Recognition' (FineReader Pro 9). While accuracy has greatly improved in font type and font size recognition, OCR technology also makes intelligent use of hardware technologies, such as 'multi core parallel processors' for speeding up the OCR process. The trends in OCR technology show that significant developments can be expected in the future.

Storage

In order to provide longevity of digital files, they need to be stored in a reliable, controlled environment (White). Master files should be stored on high quality, industry standard devices, such as CD-R, DVD, or other contemporary reliable media. Backups of master files must be created regularly and stored off-site in a secure location.

A ***RAID*** (Redundant Array of Inexpensive or Independent Disks) consists of a number of drives which collectively act as a single storage system. The production of digital material requires sufficient hard disk capacity to store files at various stages of the preservation process. It may be appropriate to consider a RAID solution if the production environment large.

In 2008, INIS purchased a THECUS N5200B PRO, 5x3,5" SATA Raid. The equipment has 5 disks of 1 TB each and has been configured as local network data storage.

Backups of master files must be created regularly and stored away from the original source in a secure location on a routine basis.

Since the beginning of the system until 1997, INIS converted all full texts of NCL from paper into microfiche for safer long-term storage. In 1997, a complete collection of NCL on microfiche, representing the intellectual knowledge and information of INIS Member States, was donated to the Central Library of Physics of the University of Vienna, which acts as a secure 'off-site' storage. As the microfiche collection is being converted to PDF, all digitized resources are also backed-up in PDF at the Central Library of Physics, which is situated less than 5 kilometers from the IAEA.

Preservation planning

In order to ensure that the contents of a digital archive remain a readable and usable information resource for the future, digital files should regularly be refreshed to new media (Hedstrom & Montgomery, 1998). This can be achieved by using different processes. The process of copying files from one storage medium to another medium of the same kind is called ***refreshing*****.** This targets media

obsolescence. After media refreshing, a verification procedure should be applied (e.g. checksum) to ensure the authenticity and integrity of the files.
Another process is ***migration***, i.e. transferring digital information from one hardware and software setting to another or from one computer generation to subsequent generations. Migration can also be format-based, to move image files from an obsolete file format to a new format.
A third process is called ***emulation*** which involves the re-creation of the technical environment required to view and use a digital resource. This is achieved by maintaining information about the hardware and software requirements so that the system can be reengineered. Due to its cost and the time required for proper emulation, this process is not often used.
At present, INIS applies the technique of **refreshing** the digital files by copying the collection to a new storage media, e.g. CD to DVD, Blu-ray Disc, etc. At the time of implementation of PDF/A as long-term archival format, the **migration** technique will be applied to all INIS digital files.

Metadata creation
Metadata plays a key role in describing, processing, managing, tracking, accessing and preserving digital resources. According to NISO[vi] (2004), metadata is key to ensuring that resources will survive and continue to be accessible into the future.
There are different types of metadata that can be associated with digital resources. INIS applies comprehensive 'bibliographic metadata' which describes the intellectual content of the digitized full text and includes an extended set of bibliographic elements for identification and retrieval of the resources. When integrated to the INIS Database, digital resources are accompanied and linked to their corresponding bibliographic records (INIS, 2009). The whole process is carefully reviewed by INIS specialists and validated by computer programs and specially designed algorithms. At present, technical metadata for digital resources is generated automatically during creation of the PDF files. However, a more sophisticated approach will be considered along with the implementation of PDF/A.

Microfiche Digitization Project
The in-house digitization of the microfiche collection started in 2002 after the acquisition of a Sunrise 2000 microfiche scanner. Initially only aimed at fulfilling document delivery requests, the digitization of the full collection became an actual topic in 2003, after a release of the new INIS on-line database that supported direct access to full-texts. It was decided to outsource a substantial part of the microfiche scanning in order to support the existing in-house digitization capabilities. The contracts were issued after formal invitation to bid and the amount of microfiche scanning requests depended on the funds available. Some funding for this project was provided by the Nuclear Knowledge Management Unit (NKM) of the IAEA. Over the years, the following three contractors were engaged: EMD Austria, Prosoft Germany, and PM Dimensions Austria. It should be mentioned that good coordination and a good strategy are necessary to ensure the success of such a project. It was especially important to avoid duplication of work and to take into consideration the different digitization initiatives by the IAEA Member States. For this reason, INIS chose the country of publication as the main selection criteria, and an extensive coordination effort with the respective Member State followed each decision to digitize their part of the INIS based microfiche collection. In order to support national document and knowledge preservation efforts, the INIS Secretariat provided Member States with DVD country sets of their digitized non-conventional literature (NCL) from the microfiche. The following table gives a detailed overview of INIS microfiche digitization activities since its inception in 2003.

Table 3: Digitization of the INIS NCL Collection on Microfiche

Year	PDF	Pages	Size (GB)
2003	566	49 574	3.7
2004	19 962	1 325 217	36.5
2005	36 935	1 577 365	32.1
2006	23 163	1 367 637	33.3
2007	9 313	668 769	16.3
2008	25 675	1 228 057	29.7
2009	81 221	3 939 811	77.3
2010	33 881	1 969 110	45.9
2011	24 027	511 990	16.2
2012	20 434	843 579	40.7
Total:	275 177	13 481 109	331.8

Close to 80% of the INIS microfiche collection has been digitized since the beginning of the project. An estimated 3 million pages still need to be processed before project completion. Depending on available resources, this major project is expected to be completed within the next two years. The ultimate goal is

a complete integration of the microfiche-based NCL into the INIS Collection and online access to full-texts provided via the Google-based INIS Collection Search

The INIS Collection Search (ICS) is a free and open web access of the INIS Collection to all Internet users. Currently it holds over 3.4 million bibliographic (metadata) records and over 350 000 full-text NCL documents. This collection of documents on the peaceful uses of nuclear science and technology is now fully indexed and searchable online using Google-based technology. Around 50 000 searches and 3000 downloads are performed monthly. A link to the INIS Collection search is available from the INIS home page[vii] or directly from http://inis.iaea.org/search/.

Besides digitization of its microfiche collection of NCL, INIS is also involved with the digitization of old IAEA publications. Examples of INIS in-house efforts include the digitization of IAEA Bulletins[viii] in all available languages, accompanied by INIS bibliographic metadata; digitization of Member State's Technical Reports and Proceedings Series, done in cooperation with the IAEA Library; digitization of reports from the International Nuclear Data Committee collection (INDC); and the digitization of out-of-print IAEA publications.

Conclusion

Large digitization projects, such as the digitization of the INIS microfiche collection of historic non-conventional literature, require serious planning, substantial funds, qualified staff, awareness of standards, and well defined purpose. Lack of qualified personnel can be mitigated by outsourcing to companies specialized in large volume digitization projects. Lack of in-house knowledge about the various aspects of digitization can also be alleviated by hiring experts and consultants, but it is important to maintain consistent quality throughout all of the digitization workflow steps, at the maximum possible level. Document preparation, selection of applicable scanning techniques, type of equipment, and adherence to current standards, are all factors which will decide success or failure of any digitization effort. Digitization should not be a goal in and of itself. Its ultimate use and usefulness must always be taken into account. It is therefore imperative that meaningful and searchable metadata accompany any digitized collection with a goal of making such a repository available through appropriate online search and delivery tools. Once this is achieved, ways and means for long term preservation need to be considered and put in place in order to ensure future sustainability and availability of the digitized collection.

References

Headstrom, M. & Montgomery, S. (1998). *Digital Preservation Needs and Requirements in RLG Member Institutions*. A study commissioned by the Research Libraries Group. Mountain View, CA, USA: RLG. Retrieved from http://www.oclc.org/research/activities/past/rlg/digpresneeds/default.htm

INIS (1999). *INIS Status Report 1998. Twenty Seventh Consultative Meeting of INIS Liaison Officers.* 631-L2-TC-441.27/2. Retrieved from http://goo.gl/HQWic

INIS (2004). *INISProgress and Activity Report 2003.* L2.04.01/INIS-PAR/2003. Retrieved from http://goo.gl/nXe4p

INIS (2006). *New Guidelines for Submission of Non Conventional Literature (NCL) full text to INIS.* INIS Technical Note No. 185

INIS (2009). *Guide to Bibliographic Description* (Rev.8). June 2009, Amend. 3. IAEA-INIS-01. Retrieved from http://goo.gl/bXNPm

INIS (2010). *The International Nuclear Information System (INIS): The First Forty Years.* Prepared by C. Todeschini. Retrieved from http://goo.gl/w7hUV

INIS (2011). *INISProgress and Activity Report 2010.* Retrieved from http://goo.gl/s3yQe

NISO (2004). Understanding Metadata. Bethesda, MD, USA: NISO Press. Retrieved from www.niso.org/standards/**resources**/Understanding**Metadata**.pdf

White, J. *Guidelines for Using Electronic Records: Illinois State Archives*. Retrieved from http://www.cyberdriveillinois.com/departments/archives/records_management/electrecs.html

Endnotes

[i] The International Nuclear Information System (INIS) hosts one of the world's largest collections of published information on the peaceful uses of nuclear science and technology. INIS is operated by the International Atomic Energy Agency (IAEA) in collaboration with over 150 member states and international organizations. There are over 3.4 million bibliographic references to publications, documents, technical reports, non-copyrighted documentation and other 'grey literature' made available, as well as 350 000 full texts. INIS offers free and open online access to this unique collection of non-conventional literature through its search application (http://inis.iaea.org/search/).

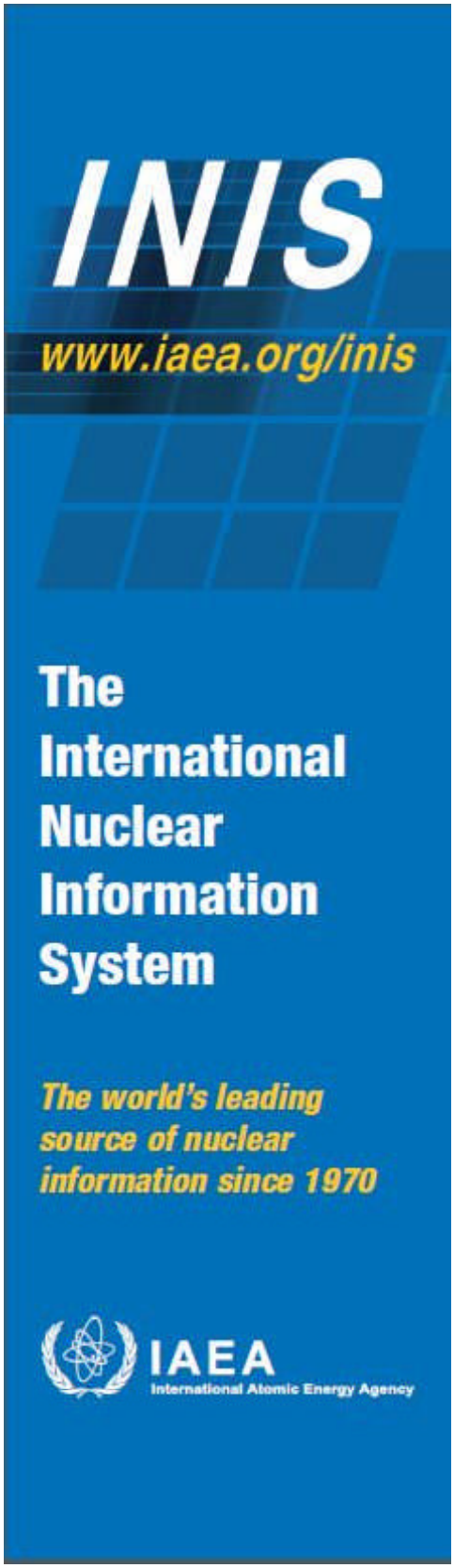
[ii] http://www.emc.com

[iii] Livelink was the first Web-based collaboration and document management system made by the OpenText. http://www.opentext.com/2/global/products/products-all/livelink-landing.htm

[iv] *PDF/A* is an ISO-standardized version of the Portable Document Format (*PDF*) specialized for the digital preservation of electronic documents. http://en.wikipedia.org/wiki/PDF/A

[v] See File compression: An introduction for raster image files. http://www.library.carleton.ca/sites/default/files/help/gis/File_Compression.pdf

[vi] National Information Standards Organization (NISO), a non-profit association accredited by the American National Standards Institute (ANSI), identifies, develops, maintains, and publishes technical standards to manage information. http://www.niso.org
[vii] http://www.iaea.org/inis
[viii] http://www.iaea.org/Publications/Magazines/Bulletin/Bull521/index.html



Korea Institute of Science and Technology Information (KISTI)

English version - http://en.kisti.re.kr/

*** Vision**

World-class information research institute creating values for customers

*** Main functions**

*** Management and service of Korean R&D reports**

KISTI exclusively manages, preserves, and serves Korean R&D reports for citizens and government officials. It provides Korean R&D reports and their information with National science & Technology Information Service (NTIS) and National Discovery for Science Leaders (NDSL).

***Contact information**

KISTI email address: hcpark@kisti.re.kr

Headquarters: Tel : +82-42-869-1004, 1234 Fax: +82-42-869-0969

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

Seon-Hee Lee and Mi Hwan Hyun
Korea Institute of Science and Technology Information, KISTI, Rep. of 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)

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

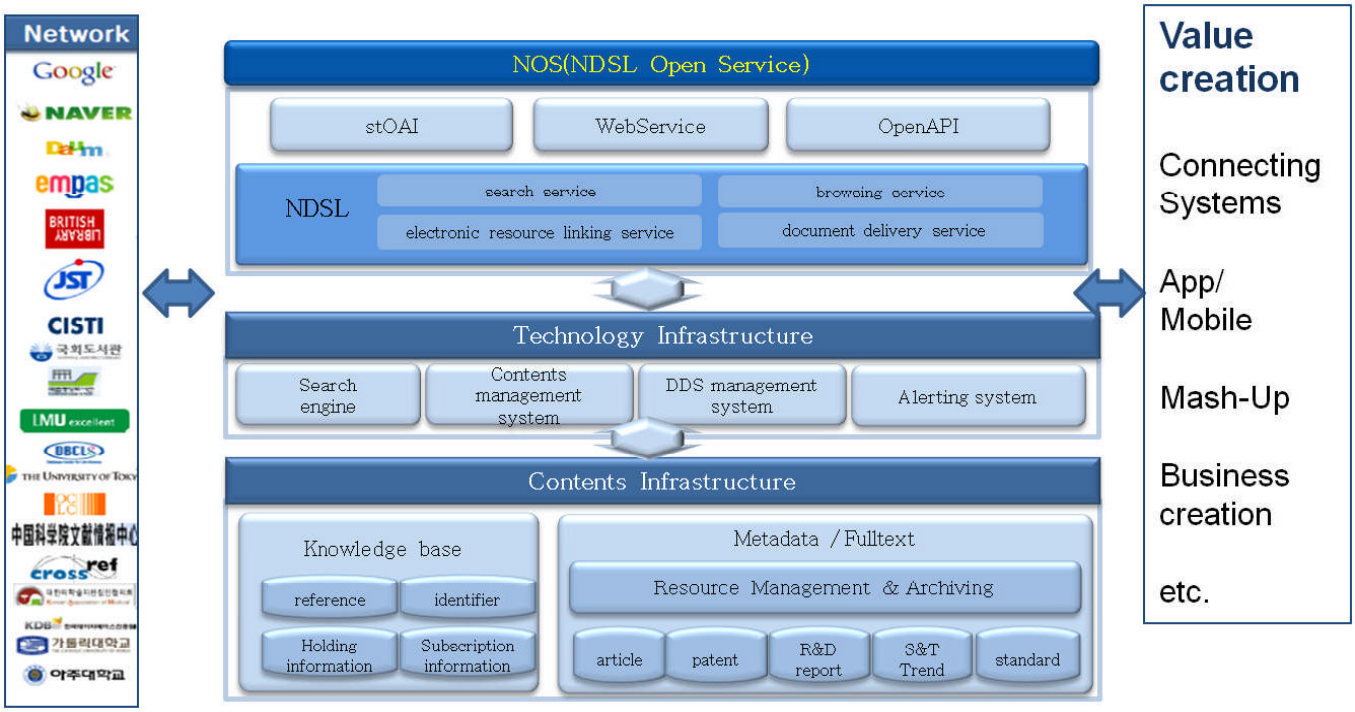
This study analyzes the effect of NOS on the national R&D technical report usage by examining the changes in the amount of usage and the number of participating organizations in NOS.

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

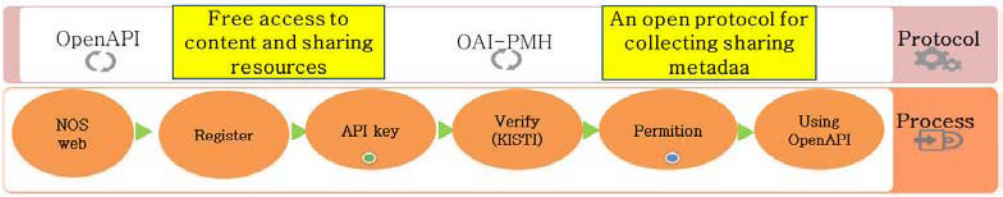
NOS utilizes contents and services in NDSL infrastructure as seen in the NOS configuration <Figure 2>. NOS provides metadata and the full-text of R&D technical reports, articles, patents, S&T trends, etc. from NDSL Resource Management & Archiving. NOS uses references, journal holding information, identifiers, 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

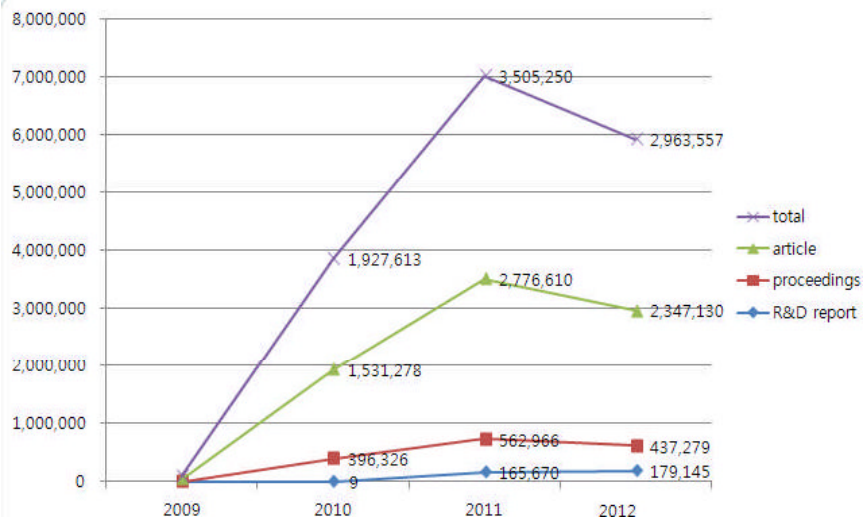
Type	stOAI	OpenAPI
Protocol	OAI-PMH	REST
Contents	domestic articles, R&D reports, S&T trends	articles, patents, R&D reports, S&T trends, standards
Services	metadata, full-text links	search, browse, DDSs(Document Delivery Service), link resolver
Data transmitting method	XML	RSS XML JSON
Authentication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

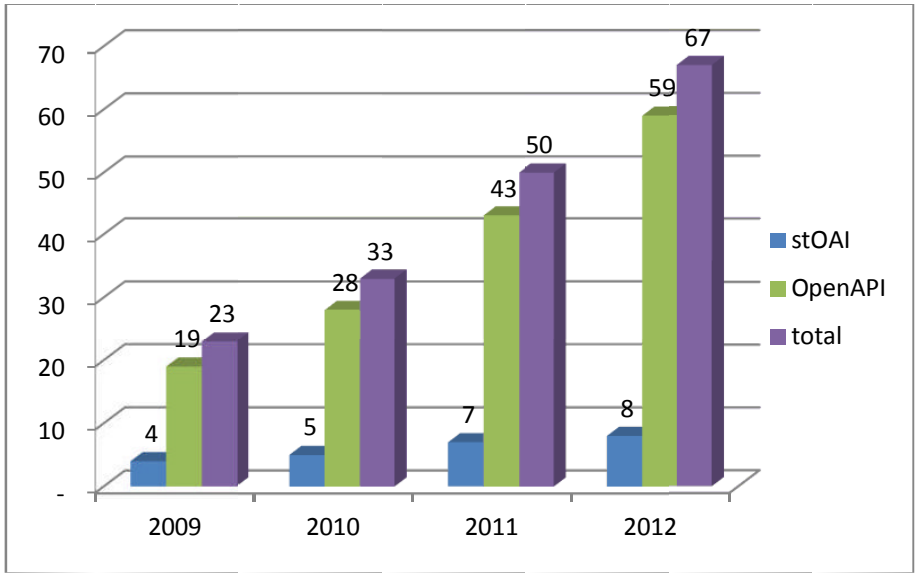
The NOS usage is demonstrated in <Figure 4>. It has dramatically increased since 2009. The total number of NDSL contents downloads was 3,505,250 in 2011 and 2,963,557 in 2012. The most popular download type of content was journal articles. The ratio of journal article downloads was 79.2% in 2012. Therefore,

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.

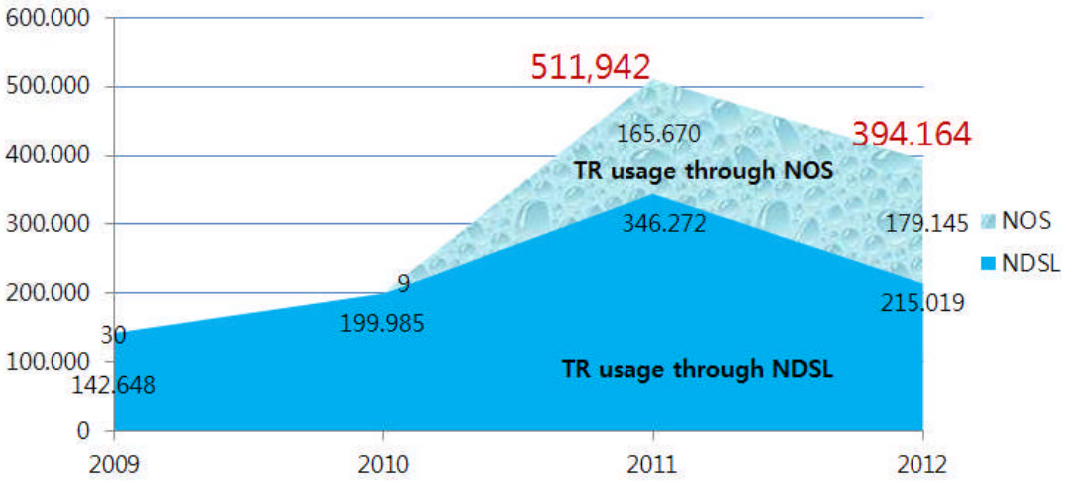


<Figure 5> NOS organizations (2009-2012)

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

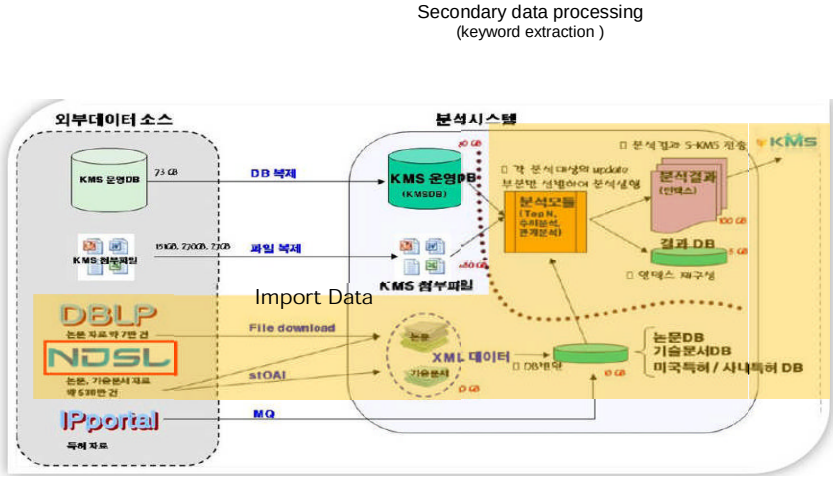
The technical reports (TR) usage through NOS in terms of full-text download has increased. The number of TR downloads in 2012 was 394,164 titles. TR downloads through NOS has increased by 83.3% since its implementation over the previous usage by the NDSL portal alone as seen in <Figure 6>.



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

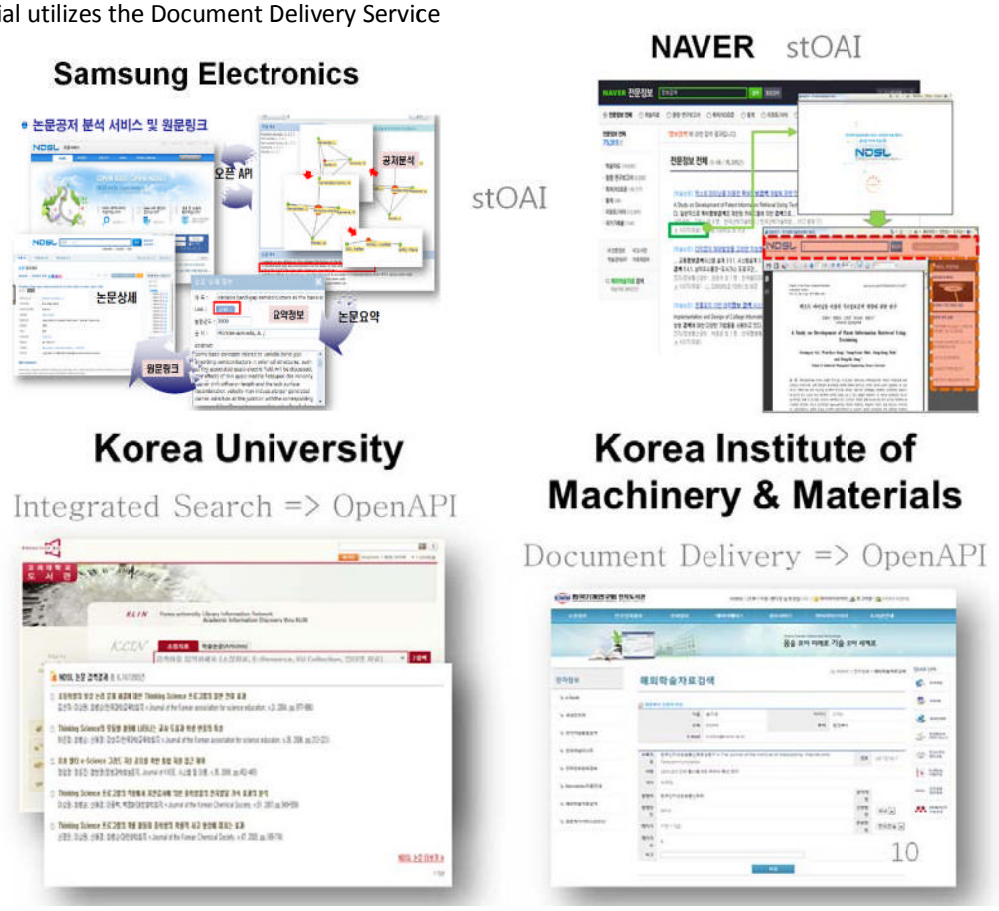
Utilizations by Organizations

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.



<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/>

Federal GL System Input Flow Analysis

Aleksandr V. Starovoitov, Yuri M. Bogdanov, Aleksandr M, Bastrykin,
Anton I. Borzykh and Leonid P. Pavlov
Centre of Information Technologies and Systems of Executive State Authorities, CITIS, Russia

Abstract
The procedures of collecting, processing and disseminating information on research and development (R&D) reports and theses in Russia are specified by the federal-level grey literature (GL) information system introduced into production use at the Centre of Information Technologies and Systems of Executive State Authorities (abbreviated in Russian as CITIS). The continuity of input document flow is secured by the federal law "On the obligatory copy of documents". The quantity of arriving documents is important from both theoretical and practical points of view: it reflects the scientific activity of academic community and an executive discipline of scientists; it determines the distribution of workload over time and the completeness of the federal collection. In the paper the system's input flow analysis is presented based on statistical records of many years. Numerical data in relative and absolute values are given.

In the Russian Federation the practices of legal deposit are spread over the most informative documents of grey literature (GL) – scientific and technical (or research and development, R&D) reports and theses. The practices are supported by the federal law "On the obligatory copy of documents" and the term "the obligatory copy of documents" is used in Russian synonymously to "legal deposit". The procedures of collecting, processing and disseminating information on research and development (R&D) reports and theses are specified by the federal-level grey literature (GL) information system introduced into production use at the Centre of Information Technologies and Systems of Executive State Authorities (abbreviated in Russian as CITIS).

The system's main information resource is the collection of two types of documents: 1) primary documents - full-text R&D reports and theses (candidate and doctoral dissertations, according to Russian tradition); 2) secondary documents - information cards (IC) containing the bibliographic description and abstracts (metadata) of the full-text documents. The information card for reports is similar to those used in the USA (NTIS Report Documentation Page, Standard Form 298) but more extended containing more than 30 fields as compared to nearly 20 fields in Form 298. Besides IC there is another type of cards arriving in the system – registration cards (RC) that are filled in by the performing organization in the beginning of every scientific work or project and usually showing if a full-text report should be prepared in the end of the research.

The essential characteristic feature of the system is that it functions in the waiting conditions when the volume of processed information, the system's load, is entirely dependent on the external factor – the volume of the input document flow.

Because of the legal deposit regulations and nation-wide coverage of the system its input flow of arriving documents is manifestly indicative of the general situation and processes in Russian science and technology.

The quantity of arriving documents is important from different points of view: it determines the distribution of workload over time and the completeness of the federal collection; it reflects the scientific activity of academic community and an executive discipline of scientists. Therefore a quantitative input flow analysis is both of theoretical and practical interest. In this paper we will consider neither content of arriving documents nor their quality standard – this would require a separate study. But in many aspects quantitative data are significant without respect to document quality: the computer power and human efforts needed to process a document do not depend on the document's scientific level as well as the collection to be complete must include all the issued documents no matter what their scientific importance is and only in case of complete collection its quality monitoring would allow to reveal that the situation in this or that science subject or research institution is up to the standard or needs improvement.

Some factors of the information flow dependence are common to the entire 45-year history of records, some are typical for a particular period of the system's functioning.

It is possible to distinguish four main factors having an influence on the input flow:

- seasonal,
- organizational,
- financial,
- legal and disciplinary.

Seasonal factor is the most evident system's input flow peculiarity: the flow is, so to say, regularly uneven. The cycles of flow peaks and low input are repeated within every year in the same months during decades and are expected to remain so as long as the traditional starting and finishing dates for scientific works exist in Russia. As a rule, the documents defining the finishing terms of scientific research and development works (contracts, agreements, plans, etc.) suppose either half-yearly (that is, in June) or yearly (in December) reports. The same is true for theses: the peaks of dissertation presenting are in November-December and in May-June. December finishing gives arrivals peak in February-March because in January there are 10-day New Year and Christmas holidays in Russia (Orthodox Christmas is on January 7th) and the peak of summer arrivals is May-July. During the peak months the input flow increases on the average two times as compared to other months and during the peak three-four months about 50% of the yearly flow arrive at the system. The system is supposed to be adaptive to processing the peak volumes of arriving documents still keeping within the time-limits of the existing regulations. The cyclic input flow changes within the year should be considered normal since they are determined by the calendar economic and financial cycles typical for all the countries.

Besides this cyclic recurrence it happens that some more general factors influence the flow from time to time, typically once in several years, resulting in a temporary downfall or, on the contrary, a rise in the flow. The influence of these factors is evident in the yearly arrivals diagrams we are considering below. Such factors may have organizational, financial or legal and disciplinary nature and must be taken into account when planning the system's workload and forecasting an expected input flow.

The organizational factor may in turn be of administrative or technological nature. The administrative reorganizations or reforms are not so rare a phenomenon and theoretically are always aimed to improve the situation, be it in economy or science. Practically this is not always so, in any case if their positive effects need a long period of time to become evident the negative effects sometimes appear much sooner. For example, the merger or division of ministries, the reorganizations of dissertation councils responsible for awarding scientific degrees often result in an unnatural drop or upsurge in the amount of arriving documents. The technological reorganizations or, rather, modernizations of the system like introducing the networking and online modes of operation always give positive effect in the rise of the incoming flow.

The financial factor is the most evident and straight one. When the funding of scientific research grows more contracts for research and development are concluded resulting in more reports prepared and presented to the system. And vice versa, the input flow decreases when the financial situation deteriorates. Mention should be made that the state, or federal budget funding still accounts for about 70% of the total science expenditures in Russia.

The last but not the least, **legal and disciplinary factors** play an important role in the flow regulation. The high level legal documents like Federal laws and Government decisions are fundamental in ensuring the completeness of the document flow and collection. The Federal law defining the legal deposit regulations has been acting since 1994 and concerns both the reports and dissertations. In respect to dissertations the existing rules of scientific degree conferment based on the Federal law and following Government decisions ensure practically 100% delivery of all the required dissertation documents (full-texts and information cards) to the system. Because the authors of dissertations cannot be given the degree without presenting the documents to the Federal GL system this disciplinary factor proves to be the strongest one and works perfectly during all the years of statistical records. Unfortunately, as far as reports are concerned the situation has not been so perfect and due to the lack of the executive discipline the completeness of the report flow varied from approximately 85% at best to 40% at worst. The sanctions against the organizations that violated the law were very weak and the motivations to observe the law in this respect were also not strong. Only recently the situation has been improved when several Government decisions were issued that connected the delivery of reports to the system with some financial privileges for the scientific organizations. For example, the organization can enjoy a reduced tax rate only if it has timely delivered all the necessary report documents to the Federal GL system.

Before getting down to the commenting on the statistical records of the yearly document arrivals it is convenient to divide the input flow history into three periods following the division into periods of the newest history of Russia. The fact is that the absolute values of the arrivals differ greatly for each of the periods not because of the factors listed above (that did work within the periods) but because of much stronger, so to say, extra-GL general factors of political and economic nature that affected the very existence of the country.

In this connection we distinguish:

- Soviet period (1967 – 1991);
- Transition period (1992 – 2000);
- Modern period (2001 – 2012).

1. Soviet period (1967 – 1991)

The period is started with the year of the system's development – 1967, but the first stable statistical records belong to 1969 – and ends in 1991, the year of the Soviet Union disintegration. Of course, technologically that was quite a different system (by the way, the first database was created on a General Electric computer made in France) but the types of arriving documents were the same. The yearly document arrivals for the Soviet period are shown in Table 1.

Table 1. Yearly document arrivals: Soviet period

Years/Doc. types	Registration cards (RC)	Information cards (IC)	R&D reports (full-text)	Dissertations (full-text)	Total (per year)[a]
1969	47500	35900	26500	21000	151900
1971	85000	72500	51170	28830	266330
1974	71670	83300	59700	28700	272070
1976	99160	105000	72300	16330	309120
1977	83300	89170	63300	16700	269170
1979	91700	106160	85500	21670	326700
1981	126258	153595	126167	21817	449654
1982	96785	122618	102897	25537	373374
1984	90021	122732	98546	24120	359539
1986	140786	177024	135817	23826	501279
1987	102746	125788	95562	24660	373416
1990	70614	97414	63602	29628	290886
1991	58968	94939	61128	27934	270903
1992	21527	28570	13417	25995	115504

First of all what strikes one's eye when looking at the data is very impressive absolute figures – several hundred thousand documents, an order of magnitude more than during next periods. Of course, the figures are given for the USSR; for the Russian Federation they were about 25% less because of the arrivals mainly from Ukraine and Belorussia and also from the Caucasian, Central Asian and Baltic Republics. The high volumes of the input flow reflect the situation when the Soviet Union was not integrated into the world economy and had to develop its own industry and a wide front of scientific research including all the branches of science and technology. In accordance with the planned economy the 100% state funding of science was regular and stable on the level of 3,5% of the gross national product. The completeness of the flow was secured by a well-organized structure of scientific and technical information departments at research organizations responsible for R&D reports preparation. The executive discipline in respect to delivering the reports to the GL system was high as the Government decisions to do so were strictly observed and controlled. The regular periods of economic development were determined by 5-year plans that also fixed the dates of scientific works starting and finishing. The peaks of arrivals are evident for the final years of 5-year plans – 1976, 1981 and, the absolute record for all times so far, 1986 (more than 500 000 documents totally).

There are more information cards than full-text reports arriving since some contracts allow that there is no full-text report in the end of research but the information card only. This is also true for the other two periods. The fact that there was more information cards than registration cards in the Soviet period can be explained by the existed practice to prepare several reports for each stage of one registered work or project.

[a] When calculating the values of this column the amount of dissertations must be doubled since each full-text document is compulsorily accompanied by the information card

2. Transition period (1992 – 2000)

This is the notorious period of the country's general crisis characteristic for the transition from the planned to market economy. Inevitably, the process badly affected all the sphere of science and technology, all kinds of scientific institutions and activities. The yearly document arrivals for the Transition period are shown in Table 2.

Table 2. Yearly document arrivals: Transition period

Years/ Doc. types	Registration cards (RC)	Information cards (IC)	R&D Reports (full-text)	Dissertations (full-text)	Total (per year)
1992	21572	28570	13417	25995	115504
1993	11438	11823	5820	14685	58451
1994	10456	8740	4760	13005	49966
1995	7665	6514	4380	9090	36739
1996	13304	11852	8260	12375	58166
2000	13879	11101	6400	24453	80286

The document arrivals data for the period reflect quite well how painful and dangerous the situation in science was. Because of the general political and economic crisis enveloped the whole country all the factors – organizational, financial, legal and disciplinary - combined in their negative influence to result in such low records. The organizational and legal disorder was amplified by the financial collapse. Instead of the Soviet 3,5% of the GNP science funding dropped to 0,5% of the GNP in 1993 and to 0,3% in 1996. In 1992-1996 the basic research expenditures of the Russian Academy of Sciences dropped 10 times [1]. And instead of total 500 000 documents in 1986 the flow sank to less than 40 000 in 1995.

Even the dissertation activities that are individual and creative in nature and less dependent on the external factors suffered greatly from the general situation of the nineties. Instead of typical for many years amount of presented dissertations around 25 000 this figure dropped twice in the years of 1993-1996 and reached the absolute minimum of 9 090 dissertations in 1995.

Thus, the main reasons for the dramatic fall of the flow during this period were the abrupt suspension of scientific research funding so that most of scientific works stopped and the decline of executive discipline in scientific organizations so that even the completed reports were not always sent to the system.

3. Modern period (2001 – 2012)

Chronologically this period began with entering the third millennium when there was a certain economic stabilization in Russia and the period still continues. The year of 2012 is just the latest year which we have completed data for. The yearly document arrivals for the Modern period are shown in Table 3.

Table 3. Yearly document arrivals: Modern period

Years/ Doc. types	Registration cards (RC)	Information cards (IC)	R&D Reports (full-text)	Dissertations (full-text)	Total (per year)
2000	13879	11101	6400	24453	80286
2004	13000	10800	6500	25460	81220
2008	14982	10536	6899	25226	82869
2010	21100	15300	11000	24700	96800
2011	31000	19600	14500	24780	114660
2012	32798	19860	14360	31229	129476

The general characteristic feature of the period is the positive dynamics in the flow of arrivals. It reflects several important measures taken to improve the situation in science using organizational, financial and legal instruments. The reformatory processes are still in progress: a radical reform of the Russian Academy of Sciences took the shape of a new law in September this year. Let us see how the taken measures are mirrored in the input flow.

First, there has been a growth in scientific research state budgeting every recent fiscal year (except the year of 2009 when because of the world economic crisis the state budget of Russia was sequestrated including the item of science expenditure). If during the transition period the funding of

science decreased then during the decade of 2002 – 2012 the state budgeting grew 10 times (from 31 billion rubles to 328 billion rubles) [2]. No doubt, this factor influenced the growth of the input flow.

Next comes the factor of introducing new technologies. In 2011 there was 50% growth in registration card arrivals (from 21100 in 2010 to 31000) and 30% growth in information card arrivals (from 15300 in 2010 to 19600) mainly because of the implementation of the online mode for filling-in the cards in screen forms and inputting them to the system in network conditions. Now 100% of cards are completed and entered online. The fact that now the amount of information cards lags behind that of registration cards indicates that something goes wrong: either several registered works were only started and not finished or if finished the reports and information cards don't arrive at the system.

During the modern period there are two noticeable advances in the arrival of full-text reports. In 2010 the amount of reports grew to 11 000 (from about 7 000 in previous years) to a great extent due to the introduction of a new paragraph in the text of a standard contract for the performing of a scientific project. The paragraph obliged all the performing organizations to deliver the report documents (full-text with IC or only IC) to the GL system. The next leap in 2011 (to 14500 reports) can be accounted for the administrative measures taken by the Russian funds for scientific and social research support that also obliged all the grant winners to register and deliver their report documents to the GL system.

The amount of dissertations remains stable around approximately 25 000 per year. The upsurge of 2012 is likely because it was announced that in 2013 there would be a radical reorganization and reduction of the Dissertation councils (which has really been carried out this year) and many people made haste to present their dissertations before reforms in 2012.

In conclusion it must be said that the input document flow to the federal GL system is indicative as a mirror showing general processes and conditions in Russian science and technology. The characteristics of the flow may be used to suggest what should be done to improve the existing situation in respect to increasing both the absolute amounts of prepared documents and the completeness of the arriving reports flow.

References

1. Commentary to the Federal Law "On science" (In Russian). Online Law Information System "Garant" (by subscription).
2. On science funding. http://www.interfax.ru/russia/txt.asp?id=273344 (In Russian). Visited Nov. 9, 2013.

PATLIB Centre: why are we grey?

Mária Harachová and Ľubomír Kucka
Slovak Centre of Scientific and Technical Information CVTI SR, Slovakia

This contribution could only be written because, according to the typology of GreyNet International, patents belong to documents of grey literature. However, this classification does not apply in all countries including Slovakia where, patents are not a category of grey literature. There are several possible reasons for that: patents are published by the Industrial Property Office of the Slovak Republic (IPO SR) and are freely accessible through the online Register of Patents.

Our institution – the Slovak Centre of Scientific and Technical Information (SCSTI) was founded as the Slovak Technical Library in 1938. One of its missions was to collect also special types of literature, such as technical standards and patents. Collection of patents begun in the 1950s and has continued ever since. Up to 1993, the collections included printed patents from the former Czechoslovakia, Germany, USA, Poland, Austria, Sweden and the former Soviet Union, as well as holdings of secondary patent literature (e.g. official journals; annual reports of the IPO SR; printed international classifications of patents, products and services; IP-related monographies and journals).

In spite of the fact that a large proportion of these documents was moved from the SCSTI to the IPO SR when it was established in 1993, the SCSTI still continues collecting printed national patent documents – patents and utility models. In addition to the Slovak patent documents, currently these holdings comprise full texts of European patents since 1993 (granted and published by the European Patent Office - EPO) on CDs or DVDs, USA patent documents since 1995 (granted and published by the United States Patent and Trademark Office - USPTO) and annotations of Japanese patent applications since 1989 (published by the Japan Patent Office - JPO). The total number of all the patent documents printed, stored on CDs or DVDs in the collection of the SCSTI amounts to 5,642,200 which corresponds to approximately 259,520 physical units. All the documents are located in the reading room of special literature and in its surrounding premises on the 4th floor in our institution. Our staff provides assistance to users, e.g. in finding patents, which are available in the reading room.

All patent documents are accessible also online in patent databases. At present, patent documents are browsed preferably via the Internet; the printed and CD/DVD documents are not required by users any longer. In the reading room, there are four PCs and users are provided with free access to national and regional patent databases: national registers, Espacenet (EPO), Patentscope (World Intellectual Property Organisation - WIPO), Depatisnet – a German database and EAPATIS - Eurasian Patent Information System, created by the Eurasian Patent Office (EAPO). A commercial licenced database Derwent Innovations Index (a part of the Web of Knowledge platform) is accessible to all users registered in our institution also via remote access. Global Patent Index (GPI), which is a part of the patent information services for experts and is available on the EPO website, is used by the PATLIB staff only.

The effort of the European Patent Organisation to increase public awareness of industrial property and its protection and make patent information available to the public has led to creation of PATLIB Centres which form a network of patent information centres. They were established in all the EPO member states by means of national patent offices. Currently, the PATLIB network consists of 340 Centres which provide users with local access to patent information and related issues. The abbreviation PATLIB, which stands for PATent LIBrary, explains their role. The staff in the Centres is skilled and provides practical assistance in the field of intellectual property rights. Its activities comprise a number of services; the staff provides intellectual property (IP) documentation, copy and document delivery services, performs searches in IP documentation, monitors technology and competitors' activities, carries out technology and competitor trend analyses, organises and is actively involved in database and search training programmes, patent clinics, consulting services, etc.

In Slovakia, five PATLIB Centres are active in the Industrial Property Office of the Slovak Republic in Banská Bystrica, the Slovak National Library in Martin, the State Scientific Library in Banská Bystrica, the State Scientific Library in Košice and in the Slovak Centre of Scientific and Technical Information in Bratislava.

The PATLIB Centre in the SCSTI was established in 2003. It provides the majority of the above services, but it focuses on preparation of patent searches supplemented with comments on search results in particular. Qualified search experts search the vast amount of information stored in patents and advise

clients so as they understand the patent information acquired. Thanks to development in IT, searches are performed via the Internet in patent databases which renders this work convenient and effective.

In October 2009, the Administrative Board of the European Patent Organisation approved the document "Cooperation programme of re-orientation of PATLIB Centres". This programme begun as a pilot project involving a limited number of PATLIBs in November 2010. Its main objective was to re-orient PATLIB Centres so as to become the centres which provide advanced, sophisticated services. The further objectives were to prepare Centres for providing patent-related innovation support services for clients in order to support their business activities, to provide these services in the pre-filing and commercialisation phases of inventions, to motivate centres to establish a network of innovation-support institutions and create direct contact with experts. Participants in the pilot project were 17 PATLIB Centres from 11 countries (the Czech Republic, Estonia, Finland, France, Italy, Latvia, Poland, Romania, Slovakia, Spain and Turkey. The PATLIB Centre at the SCSTI is the only from the Slovak Centres which was interested in participating in this pilot project. The Industrial Property Office of the Slovak Republic accepted this decision and the EPO approved our nomination.

The endeavour of the coordinator of the pilot project was to achieve a higher professionalisation of services provided by PATLIBs. For this reason the first phase of the project was aimed at improving search services. In the second phase, supplementary services, such as commercialisation support, patent strategy development, assessing economic value of IP, networking contacts, became a focal point. The last phase is focused on business development and the sustainability of PATLIBs.

Our PATLIB team, as well as, the staff from the other Centres benefit from the pilot project; all the participants have been trained regularly and coached by professionals from the EPO and also by experienced experts. This system has created conditions for the staff to be able to analyse and understand all the details of patent documents and carry out more precise patent searches. The PATLIB Centres involved are in the position to introduce new advanced, sophisticated services and the staff is well-prepared to operate the PATLIB Centres successfully. Ultimately, increased activities of these Centres help them enhance awarness of the existence and activities of the Centres in the professional as well as the general public.

Based on the SCSTI team's experience it can be concluded that improvement the PATLIB staff's qualifications, competences and skills has contributed to more considerable dissemination of patent information and increased the awareness of IP rights and patent system. Over three years of the implementation of the project, our PATLIB team, in addition to being involved in other activities, have organised 13 training seminars for our users. The participants in these training seminars gained the basics on patents, trademarks and designs and other IP rights; national, European and international patent databases and principles of searching in databases; relevant information resources; protection of ideas and inventions, i. e. intellectual property in general.

The pilot project has also helped intensify our promotional activities and in such a way enhance reputation and visibility of our PATLIB Centre. The number of clients interested in patent searches, aiming to protect their IP, has increased significantly. And this is a very important achievement resulting from a successful implementation of the pilot project. The larger the number of outcomes of scientific research and inventions protected, the larger the number of patents potentially granted; this is the fact that underlines the importance of grey literature.

Contribution to the improvement of dissemination of grey literature - JAEA Library's efforts for collecting, organizing and disseminating information on nuclear accidents

Kiyoshi Ikeda, Takeshi Ohshima, Mayuki Gonda, Shun Nagaya, Misa Hayakawa, Yukinobu Mineo, Minoru Yonezawa, and Keizo Itabashi
Japan Atomic Energy Agency, JAEA, Japan

Abstract
In March 2011, the Fukushima Daiichi Nuclear Power Station Accident occurred in Japan. After the accident, parts of information on the accident has been distributed as grey literature and they cause some issues in terms of permanent accessibility to the information. This paper introduces the activity of the Japan Atomic Energy Agency Library as an example of efforts to improve access of Internet information using the DSpace. Consequently, we contributed to the improvement of the dissemination of grey literature at the following three points; (1)to ensure permanent access to Internet information by cooperating with the National Diet Library's Web archiving project, (2) to develop standardized metadata schema and classification system, (3)to develop prototype system using DSpace and compile about 36,000 metadata.

1. Introduction

In March 2011, the Great East Earthquake occurred off the northeast coast of Japan. The Fukushima Daiichi Nuclear Power Station of Tokyo Electric Power Company (hereinafter referred to as "TEPCO") Accident (hereinafter referred to as "the Fukushima Accident") occurred. After the Fukushima Accident, Japanese governments, research institutes and TEPCO started releasing information on the accident by a various kind of information media, such as literatures like books, articles, oral presentation information, proceedings, technical reports and Internet information.

Because it is very important for us to preserve experience and know-how obtained from the Fukushima Accident, Japan Atomic Energy Agency (hereinafter referred to as "JAEA") Library has started activities to collect and disseminate such the Fukushima Accident information above. However, parts of information on the Fukushima Accident have been distributed as grey literature and they cause some issues in terms of permanent accessibility to the information.

It may be true that Internet information is not grey literature because it is usually open to the public. However, normally Internet information does not have permanent accessibility, there is a lack of bibliographical control, and some of the websites have been inaccessible and disappeared. Thus, on the contrary, it should be mentioned Internet information is Grey literature. In addition, it is difficult to get oral presentation information of latest research results at the local meetings related to the Fukushima Accident, especially held in Japan, because proceedings of oral presentation is not always published as literature in print and electronic formats, and only participants of the meetings could share information.

In this paper, the authors introduce how they contribute to the improvement of the dissemination of grey literature, i.e. internet information and oral presentation information on the Fukushima Accident by developing the archive system of DSpace, creating the metadata and organizing them with classification system.

2. Current status of JAEA Library activities

2.1 Outline of JAEA Library

JAEA is a comprehensive research and development organization that aims at the practical use of nuclear energy and is involved in basic and applied research through the establishment of nuclear fuel cycle.

Engaged in JAEA, its Library is one of the largest nuclear information centers in Japan, and has supported the scientific research and development activities throughout Japan. JAEA Library collects and provides information in the fields of nuclear science. JAEA Library also publishes JAEA Reports originated from JAEA's R&D results and distributes them widely. Part of the R&D results includes their full-texts being available on the JAEA Library website. Another main role of the JAEA Library is the International Nuclear Information System (hereinafter referred to as "INIS") National Center for Japan. INIS is the open access database for published scientific literature on the peaceful uses of nuclear science and technology.

Since the Fukushima Accident, many information requests on the accident have been add ressed to the JAEA library. In order to meet the requests, we began collecting, compiling and distributing Fukushima Accident reference information from the special website immediately after the accident. JAEA Library has also sent INIS input records of literature related to the Fukushima Accident to the International Atomic Energy Agency (hereinafter referred to as "IAEA").

2.2 Special website of the Fukushima Accident's reference information

JAEA Library has been disseminating information related to Fukushima Accident via the special website since April 2011 after the occurrence of the Fukushima Accident (Fig.1).

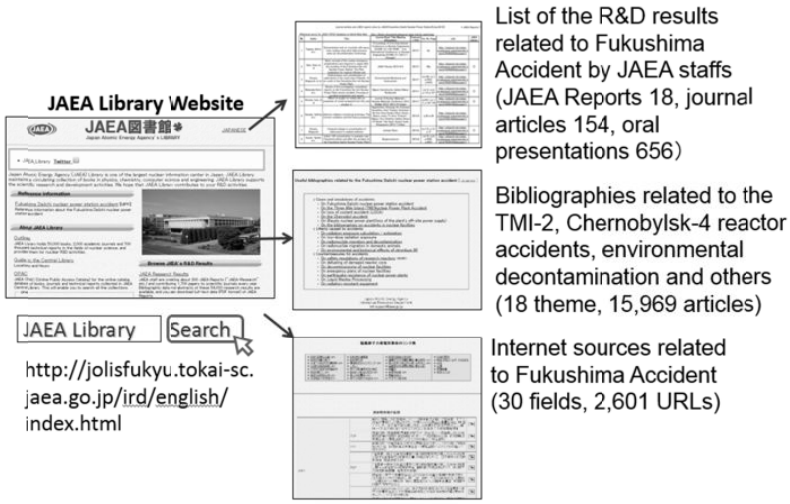


Fig.1 Special website of JAEA Library related to Fukushima Accident

The special website consists of the following contents:

- Lists of the JAEA research and development results,
- Useful bibliographies (radiocesium dispersion simulation, health effects, contamination countermeasures etc),
- Internet sources.

The list of the JAEA R&D results provides journals or conferences paper and JAEA Reports and oral presentations related to the Fukushima Accident by JAEA staffs. JAEA Reports have been made downloadable in full text freely as of today via the Internet.

The 'bibliographies' provides eighteen themes of bibliographic data on list related to the Fukushima Accident, such as Three Mile Island accident, Chernobyl accident, radioactive liquid waste processing, radionuclide migration, decontamination etc. It provides about sixteen thousand articles of bibliographic data in total.

In Internet sources, we collect about 2,000 documents on the Internet related to the Fukushima Accident and categorize them into 30 themes. Types of the collected documents include full -text, html, figure, photo, and movie data related to the Fukushima Accident. They include, for example, reports compiled by the Japanese Government, Diet, NPO, etc.[3][4][5][6]. Furthermore, raw data of technical information such as plant parameters and radiation monitoring. Our libra ry staff check all of the links to the Internet sources quarterly because links don't provide permanent access and they are sometimes changed or deleted.

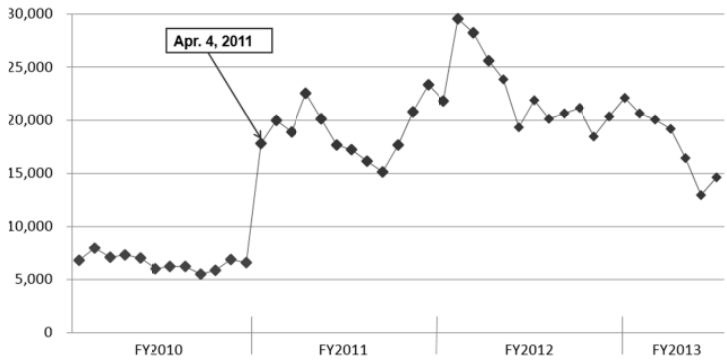


Fig.2 Number of Access of JAEA Library website

Fig. 2 shows the total number of access of JAEA Library website. It is obvious that number of access has rapidly increased after the release of the special website in April 2011. There have been about 15,000 accesses per month to the Fukushima Accident related information of JAEA Library website.

3. Development of the Fukushima Accident Archive System

Our special website originally doesn't provide search funcitionality, it is not possible to search entire information on the website and nor is permanent access usually ensured regarding internet information. Therefore, we decided to develop a system which has search functionality and selected DSpace for the system because DSpace is a commonly used as an institutional repository system, supporting OAI -PMH and capable of processing standardized metadata (Fig.3).

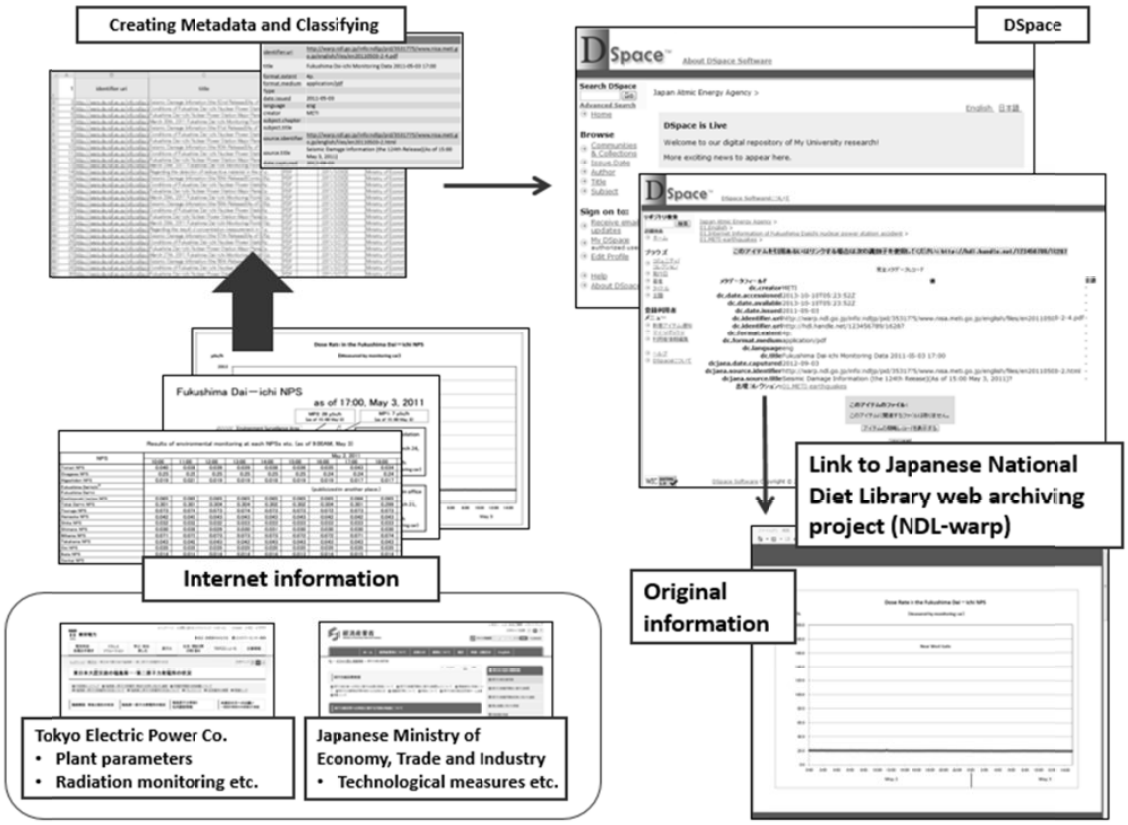


Fig. 3 Fukushima Accident Archive System to disseminate of Internet Information with the metadata and Classification System

To achive the dissemination of Internet information on the Fukushima Accident and the contribution to the improvement of dissemination of grey literature, there are challenging Issues: (1) To ensure accessibility of Internet information, (2) To organize information in a systematic manner.

3.1 Ensuring accessibility to Internet information.

To solve the first issue, we have started to cooperate with the National Diet Library (hereinafter referred to as "NDL"). The NDL is the only national library in Japan and it acquires all materials published in Japan and preserves them as the only depository library in Japan. NDL compiles catalogs of these publications in a database or other format, and with these collections provides library services.

The NDL has been providing the Web Archiving Project (WARP) since 2002. In WARP the NDL has collected inforomation in the form of website of the following: the government, the Diet, the courts, local governments, independent administrative organizations, universities and cultural and international events held in Japan. A part of these archived websites is provided on the Internet. We have con nected our metadata directly with the appropriate archived information in NDL's WARP via hyperlink. As a result of cooperation between JAEA and NDL, various kinds of Internet information will be available and permanent access to these internet ifnromation will be ensured. We have already extracted and created about 36,000 metadata from the Ministry of Economy, Trade and Industry (hereinafter referred to as "METI") and TEPCO websites.

3.2 Organization of information in a systematic manner

3.2.1 Metadate Scheme

To solve the second issue, we have needed to establish bibliographic control rule, i.e. metadata format for Internet information. We have chosen Simple Dublin Core(Simple DC) as our metadata format(Fig.4). In developing metadata of Internet information, we have needed some considerations on bibliographic control rule.

Element	Value (ex)
identifier.uri	http://warp.ndl.go.jp/info:ndljp/pid/3531775/www.nisa.meti.go.jp/english/files/en20110503-2-4.pdf
title	Fukushima Dai-ichi Monitoring Data 2011-05-03 17:00
format.extent	4p.
format.medium	application/pdf
Type	
date.issued	2011-05-03
language	eng
creator	METI
subject.chapter	
subject.title	
source.identifier	http://warp.ndl.go.jp/info:ndljp/pid/3531775/www.nisa.meti.go.jp/english/files/en20110503-2.html
source.title	Seismic Damage Information (the 124th Release)(As of 15:00 May 3, 2011)
date.captured	2012-09-03

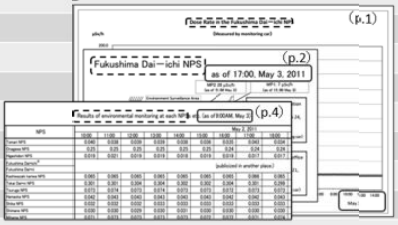


Fig.4 Metadata of Internet information related to Fukushima Accident as an example

Concerning date element, literature has only one date element i.e. publication date. But Internet information has more than one date element, such as release date and time, update date and time etc. Then we have to decide one date element for Internet information.

Another instance is the title element. Usually pictures themselves have no titles, on the other hand, some materials have more than one title. In such case, it is difficult for users to know the content of the material without appropriate metadata. Eventually, we complement the title by adding time, etc and add URL of upper level of the website page which has the link to the original information as source information on "source.identifier" and "source.title".

3.2.2 Classification System

Classification System is very useful to organize collected informatin as well as to let users access necessary information easily. So, creating Classification System is important work for us.

	Classification -broad categories-	Report of Japanese Government to the IAEA Ministerial Conference on Nuclear Safety - The Accident at TEPCO's Fukushima Nuclear Power Stations - (June 2011)	The official report of Executive summary The Fukushima Nuclear Accident Independent Investigation Commission (July 2011)	Investigation Committee on the Accident at Fukushima Nuclear (July 2011)	Independent Investigation Commission on the Fukushima Daiichi Nuclear Accident (February 2011)
1	current situation in Japan Before the Accident	I. Introduction	I. Was the accident preventable?	I. Introduction	Part.1 The damage and accident responses at the Fukushima Dai-ichi nuclear plants
2	Disaster damage in Japan from the Tohoku Distrist - Off the Pacific Ocean Earthquake and Resulting Tsunamis	II. Situation regarding Nuclear Safety Regulations and Other Regulatory Frameworks in Japan Before the Accident	1.1 Essential lack of robustness against earthquakes	1. Overview of the Investigation Committee	Part.2 The emergency responses taken by the government ministries, local government and first responders; crisis management by Prime Minister's Office, and the reality of chaotic evacuation
3	Occurrence and progression of the accident at the Fukushima NPSs	1. Legislative and regulatory framework for nuclear safety	1.1.1 Overview of the Fukushima Daiichi NuclearPower Plant	2. Activities of the Investigation Committee	Part.3 The historical, consensual and structural factors that led to the negligence of safety improvement
4	Responses to the nuclear emergency	2. Mechanism for nuclear emergency responses	1.1.2 Changes in the seismic safety evaluation of the Fukushima Daiichi Nuclear Power Plant	3. The Relationship between the Final Report and Interim Report	Part.4 Global contexts: international nuclear safety regime, nuclear security and US-Japan relations
5	Discharge of Radioactive Materials to the Environment	III. Disaster damage in Japan from the Tohoku Distrist - Off the Pacific Ocean Earthquake and Resulting Tsunamis	1.1.3 Seismic vulnerability of the Fukushima Daiichi Nuclear Power Plant when it was first constructed	II. The Damage and Accident Responses at the Fukushima Dai-ichi NPS and the Fukushima Dai-ni NPS	※each table of contents is Japanese only.
6	Situation regarding Radiation Exposure	1. Damage by the earthquake and tsunami in Japan	1.1.4 Establishment of the "Regulatory Guide for Reviewing Seismic Design" and seismic safety backcheck	1. The Damage at the Major Systems and Facilities of Units 1 to 3 of the Fukushima Dai-ichi NPS	
7	Cooperation with th International Community	2. Damage caused by earthquake and tsunami hitting Fukushima NPSs	1.1.5 Fatal flaws in the seismic backchecks against the revised guide for reviewing seismic design	(1) Introduction	
8	Communication regarding the Accident	3. Seismic and tsunami damage to other NPSs	1.1.6 Did obsolescence have an impact on the occurrence of the accident?	(2) Major instrumentation mechanisms and related discussions	
9	Further efforts to settle the accident in the future	4. Assessment of earthquake and tsunami damage	1.2 Conclusion regarding the seismic vulnerability of the Fukushima Daiichi Nuclear Power Plant prior to the accidents	(3) Analyses on various phenomena concerning severe accidents	
10	Response at other NPSs	IV. Occurrence and Development of Accidents at the Fukushima Nuclear Power Stations	1.2 Tsunami risk recognized but lacked countermeasures	(4) Discussions on the damage to major Systems and Facilities of Unit 1	
11	Lessons Learned from the Accident Thus Far	1. Outline of Fukushima Nuclear Power Stations	1.2.1 Changes in tsunami postulates and damage predictions over time	(5) Discussions on the damage to the major Systems and Facilities of Unit 2	
		2. Safety Assurance and Other Situations in Fukushima NPSs	1.2.2 Making light of the vulnerability that leads to loss of all ac power sources and core damage due to tsunamis	(6) Discussions on the damage to the major Systems and Facilities of Unit 3	
		3. Condition of the Fukushima NPSs before the earthquake	1.2.3 Problems around the review of tsunami height	2. Examination of a Hydrogen Gas Explosion	
		4. Occurrence and progression of the accident at the Fukushima NPSs	1.2.4 Lack of transparency surrounding NISA and JNES	(1) Types of explosions	
		5. Situation of Each Unit etc. at Fukushima NPS	1.2.5 Arbitrary use of probability theory	(2) Characteristics of hydrogen gas explosions	
		6. Situation at Other Nuclear Power Stations	1.3 Severe accident countermeasures disregarded international standards	(3) The explosion at the Unit 1 R/B	
		7. Evaluation of accident consequences	1.3.1 Ineffectiveness exposed by the nuclear accident	(4) The conditions at the Unit 2 R/B and S/C	
		V. Response to the nuclear emergency	1.3.2 Passive reviews as a result of collusion between operators and regulatory authorities	(5) The explosion at the Unit 3 R/B	
		1. Emergency response after the accident occurred	1.3.3 Narrow coverage of severe accident countermeasures and delayed implementations in Japan	(6) Explosion at the Unit 4 R/B	
		2. Implementation of environmental monitoring	2. Escalation of the accident	3. Response to the Accidents at Units 5 and 6 of the Fukushima Dai-ichi NPS	
		3. Measures for agricultural food stuffs and drinking water, etc.	2.1 How the accident developed and an overall review	(1) Overview of the response to the accidents at Units 5 and 6 of the Fukushima Dai-ichi NPS	
		4. Measures for additional protected areas	2.1.1 Further understanding of the accident	(2) Overview of Units 5 and 6 at the Fukushima Dai-ichi NPS	
		5. Assessment of nuclear emergency response	2.1.2 Key damage and the impact of the earthquake and tsunami	(3) Extent of the damage to Units 5 and 6 at the Fukushima Dai-ichi NPS	
		VI. Discharge of Radioactive Materials to the Environment	2.1.3 Progression of the reactor accidents	(4) Situation from the moment of the earthquake until the arrival of the tsunami (between 14:46 and 15:35 on March 11)	
		1. Evaluation of the amount of radioactive materials discharged to the air	2.1.4 Process of radiation release based on the parameters of the reactor	(5) Situation concerning reactor depressurization and water injection into the reactor	
		2. Evaluation on the amount of radioactive materials discharged to the sea	2.1.5 Efforts to prevent accidents and the accident risk at other nuclear power plants	(6) Status of the SFP up until the restoration of the RHR system and the response to this	
		VII. Situation regarding Radiation Exposure	2.1.6 Discussion	(7) Situation from the restoration of the RHR system until the cold shutdown	
		1. Situation of radiation exposure concerning radioworkers and other related workerworkers	2.2 Analyses and discussions on some issues	4. Restoration of External Power to the Fukushima Dai-ichi NPS	
		2. Response to radiation exposure of residents in the vicinity and the overall situation	2.2.1 Seismic ground motion at the Fukushima Daiichi nuclear Power Plant due to the Great East Japan earthquake	(1) Overview of the external power installations to the Fukushima Dai-ichi NPS	
		3. Evaluation of the status of radiation exposure	2.2.2 Possibility of damage to important devices due to the earthquake motion	(2) The damage to the external power installations to the Fukushima Dai-ichi NPS	

Fig.5 Comparison of the contents of the Fukushima Accident investigation reports

We have first considered developing Classification System by referring and comparing the contents of the accident investigation reports [3][4][5][6] (Fig.5) because these reports discuss diversity of

perspectives about the Fukushima accident. We think it is best to extract items that are common to those reports and to organize them.

Nuclear Knowledge Management Section of the IAEA is currently developing Nuclear Accident Knowledge Taxonomy for organization of data, information and knowledge from major nuclear accidents. This Taxonomy is planned to be used to disseminate those to IAEA member states. Then we have also started to consider to refer to IAEA's Taonomy for our Classification System because IAEA's Taxonomy is very useful in the point that it specializes in accident for preserving and organizing knowledge of experience and know-how obtained from major nuclear accidents like Three Mile Island Accident and Chernobyl Accident.

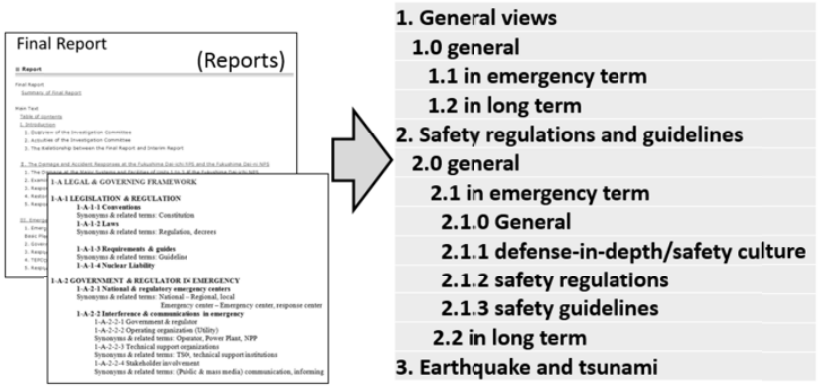


Fig.6 Development of the original classification system

But in the case of the Fukushima Accident, some situations are unique and are different from the past severe accidents like Chernobyl and Three Mile Island. For example, the Fukushima Accident has just moved from the emergency phase which is unlike the other two accidents. Thus, if we make use of IAEA's Taxonomy we need to add more categories on uniq ue situations and long term management phase, etc to the IAEA's Taxonomy. For example, categories on earthquake, tsunami, and the removal of spent fuel and debris in nuclear reactor. We will give feedback to the IAEA Taxonomy, and we hope that our activity will result in the drastic improvement of convenience for them.

We are currently developing the original Classification System by a combination of these two classifications (Fig.6).

3.3 Oral presentation information

Many conferences, symposiums on Fukushima Accident have been hold after the accident and latest research results have been presented at the conferences. It is difficult to obtain and access such oral presentation information even if oral presentation is published in paper or in an electronic form because it is usually distributed only to the participants of the conference. Thus, oral presentation information is Grey Literature. So, we try to improve access to oral presentation information because it contains current research results which are cutting-edge and valuable information on the Fukushima accident.

Thus, we have created a metadata of about 2,000 oral presentations information on the Fukushima Accident in five conferences in Japan, such as annual meeting of the Atomic Energy Society o f Japan. We have adopted Simple DC for oral presentation information as same as the Internet information.

Element	Value (ex)
identifier.uri	http://library-documents.jaea.go.jp/opac/br_search_detail.asp?id=589152&syubetsu=1
title	Risk communication practice after the Fukushima nuclear accident
creator	Ayame, Junko(Japan Atomic Energy Agency); Sugiyama, Kenji(Japan Atomic Energy Agency); …
source.identifier	http://www.aesj.or.jp/meeting/index-e.html
source.title	General Issues, 2013 Fall Meeting of the Atomic Energy Society of Japan,3-5 September 2013. Hachinohe Inst. of Technology, Aomori

Fig.7 Metadata of Oral presentation information related to Fukushima Accident as an example

As for oral presentation information we have input JAEA library holding information into the "identifier.uri" field of the metadata to ensure access to the full text of the preprint appropriately. In

addition, we have input date of the conference into the source field so that users can know conference name, session name, etc. on "source.identifier" and "source title" (Fig.7).

4. Summary

We contributed to the improvement of the dissemination of grey literature at the following ongoing three points.

1. To ensure permanent access to Internet information by cooperating with the National Diet Library's Web archiving project WARP.
2. To develop standardized metadata schema and classification system for the Fukushima Accident Internet information as well as oral presentation information. We contribute to the organization of grey literature.
3. To develop prototype system using DSpace and compile about 36,000 metadata and put them on the system. We also add information of 2,000 oral presentation information after we review and check them.

We will open to the public the Fukushima Archive using the DSpace next year.

5. Future plan

We would like to undertake the further development of the Fukushima Accident Information Archive.
We have created metadata from METI and TEPCO's websites. It is only a limited scope of a huge amount of information inside and outside Japan. So, we are going to collect, organize and disseminate as much metadata as possible through cooperation with related organizations.
JAEA library and NDL have already had cooperation on using archived information in NDL's WARP via hyperlink. In addition, we will provide metadata to NDL Great East Earthquake Archive "HINAGIKU" which is one of the outcomes of WARP, so that NDL could make use of them.
We will also provide our metadata to INIS according to the INIS rule in order to enrich contents of INIS collection because INIS disseminates nuclear literature information as well (Fig.8).
We believe that cooperation with relevant organizations is essential for the success and growth of the Fukushima Archive. Our work has just started and the collection of Internet Information is limitted, JAEA Library is going to collect, organize and disseminate as much information as possible for the future.

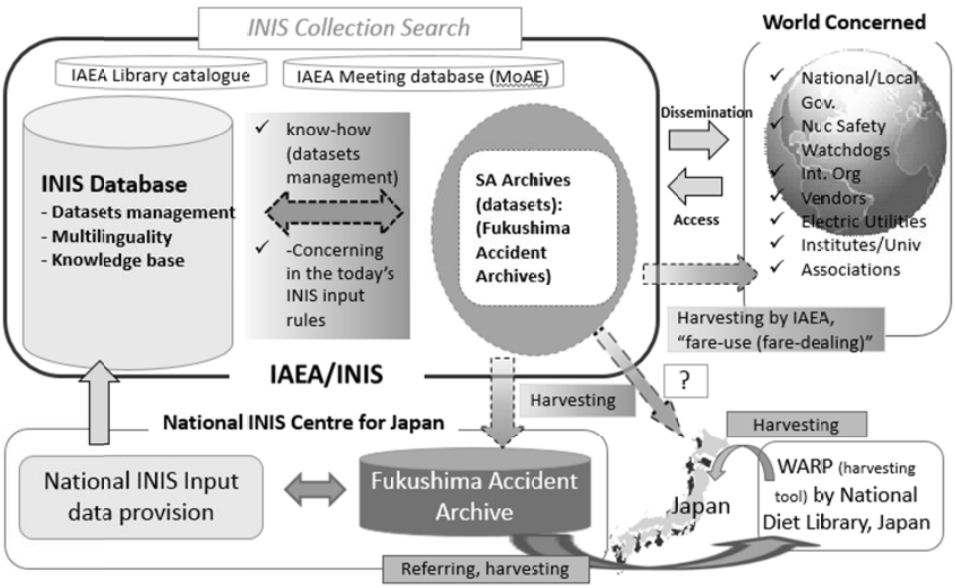


Fig.8 Cooperation with INIS

References

[1] Ikeda, K., Nakajima, H. and Yonezawa, M. (2012) 'Introduction of Useful Bibliographies and Resources related to the Fukushima Daiichi Nuclear Power Station Accident through the JAEA Library', ATOMOSU (ATOMO-sigma, Journal of the Atomic Energy Society of Japan), Vol.54, No.8, pp.549-553 (in Japanese).

[2] Nakajima, H., Ikeda, K., Yonezawa, M., Itabashi, K., Kiriyama, E. and Iwata, S. (2012) 'Study for development of the Fukushima Dai-ichi nuclear power station accident archive', Journal of Japan Society of Information and Knowledge, Vol.22, No.4, pp.344-353 (in Japanese), https://www.jstage.jst.go.jp/article/jsik/22/4/22_22_344/_pdf, retrieved 30 June 2013.

[3]The National Diet of Japan Fukushima Nuclear Accident Independent Investigation Commission (2012) 'Reports of the National Diet of Japan Fukushima Nuclear Accident Independent Investigation Commission', http://warp.da.ndl.go.jp/info:ndljp/pid/3856371/naiic.go.jp/en/index.html, retrieved 30 June 2013.

[4]Investigation Committee on the Accident at Fukushima Nuclear Power Stations of Tokyo Electric Power Company (2012) 'Final report of Investigation Committee on the Accident at Fukushima Nuclear Power Stations of Tokyo Electric Power company', http://www.cas.go.jp/jp/seisaku/icanps/eng/final-report.html, retrieved 30 June 2013.

[5]Fukushima Genpatsu Jiko Dokuritsu Kensyo Iinkai (2012) 'Fukushima Genpatsu Jiko Dokuritsu Kensyo Iinkai Chosa Kensyo Hokokusyo' (in Eng: Investigation Report on Fukushima Daiich Nuclear Power Plants Accident by Independent Investigation Panel for Fukushima Daiich Nuclear Power Plants Accident, translated by author, K.Kunii), Rebuild Japan Initiative Foundation (in Japanese).

[6]Tokyo Electric Power Company (2012) 'Fukushima Nuclear Accident Investigation Report',http://www.tepco.co.jp/en/press/corp-com/release/2012/1205638_1870.html, retrieved 30 June 2013.

[7] NDL Great East Japan Earthquake Archive "HINAGIKU". http://kn.ndl.go.jp/node?language=en

[8] INIS Collection Search. http:// http://inis.iaea.org/search/

Guiding the Grey: The Implementation and Evaluation of a Journal Club amongst a Librarian and Clinical Practice Guideline Developers – A Cancer Care Case Study

Marcus Vaska, Knowledge Resource Service (KRS), University of Calgary**;**

Xanthoula Kostaras, Emily MacLeod, Elysa Meek, Melissa Shea-Budgell,
Guideline Utilization Resource Unit, Alberta Health Services, CancerControl Alberta;

Laurissa Watson, Population and Public Health, Research and Innovation, Alberta Health Services, Canada

Abstract

Introduction/Goal: *As a research-intensive facility located within a cancer care environment, library services provided at the Holy Cross Site closely adhere to an embedded librarian mandate, one where the librarian "actively engages in activities, possesses extensive knowledge of the researcher's work, and offers assistance above and beyond common library service expectations" (Strain, 2011).*
The Guideline Utilization Resource Unit (GURU) is composed of knowledge management specialists (KMS) and nurse facilitators (NF) who support multidisciplinary teams in developing, implementing, and evaluating provincial clinical practice guidelines (CPGs) for the diagnosis, staging, treatment and follow-up of cancer. These CPGs are evidence-based documents with consensus recommendations; they are freely available on a public website for access by practitioners and patients, and are a form of grey literature. Team members at GURU consult regularly with the librarian to ensure that the most accurate and comprehensive search strategy is used to develop CPGs. The goal of this paper is to describe the process of organizing and evaluating a journal club involving a unique collaboration between guideline developers and a librarian.
Procedure: *The journal club is comprised of three KMSs, two NFs, the GURU Manager and an embedded librarian. The group has been meeting once per month since April 2012. Each member takes turns selecting two articles related to CPG development or implementation, and is responsible for leading an informal discussion. To evaluate the usefulness of the journal club and the impact of grey literature on CPG development in Alberta, all members of the journal club (n=7) were interviewed in a focus group setting or a semi-structured interview. Transcripts of audio-recorded interviews will be qualitatively analyzed for repeated themes related to knowledge gained from, and perceived benefits of journal club meetings.*
Results: *First, we hypothesize that monthly participation in the journal club will increase members' knowledge of development, evaluation, and implementation of guidelines. Second, we believe that participants will have acquired a better understanding of the research process and how to critique current guideline research. Finally, we expect that members will report that the journal club provided the opportunity to facilitate discussions around topics that are less familiar to them. It is anticipated that this collaborative venture will further enhance the importance of grey literature and its usefulness for cancer care clinical practice guidelines.*

Introduction

Holy Cross Site

Numerous medical breakthroughs have emerged over the years, as the result of research becoming an essential component of health care (Strain, 2011). Formerly a 600-bed hospital, the Holy Cross Site was converted into a primarily applied research-centered locale. This site also served as a satellite branch of the provincial Knowledge Management Department, collectively referred to as the Knowledge Resource Service (KRS). The embedded librarian at the Holy Cross Site provides library resources and services to support evidence-informed decision making, working closely with researchers on grant applications, proposals, evaluations, and systematic reviews. As with many projects focusing on service evaluation, including the journal club discussed in this paper, the overarching goal often pertains to the issue of quality improvement, namely whether or not efforts and involvement have enhanced knowledge, led to greater efficiency, and succeeded in improving and/or perhaps changing one's everyday practice regimen.

While the term *embedded librarianship* has only received specific attention in the literature over the past few years, the tasks and duties associated with the role of an "embedded librarian" have existed in this profession for decades. As Schumaker and Talley (2009) note, an embedded librarian is

one who actively engages in relationship-building activities, possesses extensive knowledge of the researcher's work, performs complex analyses, and offers assistance above and beyond common library service expectations. This service mandate forms the basis to which the Holy Cross librarian adheres on a daily basis when meeting with clients. Aside from participating in monthly journal club meetings, the Holy Cross Site librarian has provided consultation on search strategies used to support the development of clinical practice guidelines, assistance on exporting and formatting references, as well as facilitation of library instruction sessions and webinars. The importance of this liaising role cannot be overstated, as it provides the librarian with a comprehensive understanding of specific aspects of guideline development, leading to additional support.

Guideline Utilization Resource Unit (GURU)

Created in 2006 as a branch of the former Alberta Cancer Board (currently Cancer Control Alberta, Alberta Health Services [AHS]), GURU's primary objective is to support twelve of Alberta's tumour teams in the development of "evidence-informed clinical practice guidelines for site-specific cancers." (Guideline Utilization Resource Unit, 2013). Comprised of knowledge management specialists and nurse facilitators, GURU is involved with all aspects of guideline development, evaluation, and implementation, including surveillance, selection, synthesis, evaluation, and interpretation of evidence.

Clinical practice guidelines are a form of grey literature and are defined as "systematic statements about specific health problems intended to assist decision making" (Guideline Utilization Resource Unit, 2013). A substantial deliverable within each guideline is a thorough literature review, derived from the best evidence available to support a treatment or procedural recommendation. At the Holy Cross Site, GURU team members, particularly the knowledge management specialists, regularly consult with the librarian to ensure that the most accurate and comprehensive search strategy is used to develop the guidelines.

GURU produces between two and four guidelines annually for each tumour team. Guidelines cover most aspects of care, from diagnosis to follow-up and have been developed for most disease sites, from the most prominent cancer diagnoses (i.e., breast, lung, colorectal, and prostate cancers) to the less prominent cancers (i.e., head/neck, neuro-endocrine, and tumours with rare histologies). Upon completion, each guideline is published on the AHS website, www.albertahealthservices.ca/cancerguidelines.asp. Guidelines are periodically evaluated to determine whether practice in Alberta reflects the recommendations. Nurse facilitators then connect with the tumour teams to assist with implementation of the guidelines, especially where gaps exist between practice and the evidence.

Guidelines as a Type of Grey Literature

"Enhancing the transparency and accessibility of informally published research and information" (Australian Research Council, 2012) is a key grey literature goal. Relying on comprehensive literature reviews from the academically published "white" articles, in balance with the grey, the creation and dissemination of clinical practice guidelines in cancer care adheres closely to AHS' values of creating greater awareness in patient care in the hope that one day a cure will be found. (Alberta Health Services, 2013) Clinical practice guideline producers are able to share key information on evidence-based recommendations as soon as the guidelines are approved. As such, all of the research undertaken was fundamental in supporting the production of clinical practice guidelines, which are a key type of grey literature in cancer care.

The use of quality, evidence-based medicine is essential in guideline production; so too is use and retrieval of grey literature. Thus, the journal club, in accordance with how numerous journal clubs are run today, promotes evidence-based practice, critical appraisal, and continuing professional education (Swift, 2004, p. 67). Not only must this information be accurate and current, it must also be "reflective of scientific principles rather than tradition, habit, or outdated information" (Luby et al., 2006, p. 100). Further, in an effort to create greater awareness of the availability of this type of grey literature, guidelines are often freely and openly accessible to health practitioners for feedback and comments. Patients are also able to access the guidelines and, anecdotally, have come to the clinic ready to discuss the recommendations with their physicians. The guidelines are maintained on a regular basis; review and updates occur as needed or every one to two years.

Numerous health sciences libraries and health institutions at all levels (i.e. the Canadian Agency for Drugs and Technologies in Health [CADTH], the National Institutes of Health [NIH], and the World Health Organization [WHO]) have recognized the key role clinical practice guidelines play in grey literature-seeking pursuits.. For example, *Grey Horizon,* a grey literature current awareness blog in cancer care, has included guidelines as inclusion criteria since launching in April 2012. Numerous posts pertaining to the announcement of new clinical practice guidelines have been re-tweeted by GURU to colleagues

across various provincial tumour teams. The Knowledge Resource Service (KRS) website, launched on August 7, 2013 as a means of providing a centralized -point-of-intake for AHS staff throughout the province of Alberta, contains a separate page devoted to guidelines (http://krs.libguides.com/cancerguidelines), which is located within the Cancer Care subject guide. In addition, the webpage also connects to the Canadian Partnership Against Cancer, which offers numerous strategies and tips for appraising a high quality guideline.

Journal Clubs in Medical Disciplines/Fields

There are numerous means by which a journal club can be described. Nevertheless, a definition put forth more than a decade ago seems to have gained common, if not universal acceptance: "An educational meeting in which a group of individuals discuss current articles, providing a forum for a collective effort to keep up with the literature." (Kleinpell, 2002, p. 412). While the literature credits William Osler as founder of the first formal journal club at Montreal's McGill University in 1875, Buffington, Allen, and Farach (2008)state that a form of a journal club already existed in London nearly four decades earlier. Even though the format and structure of today's 21st century discussions may widely differ from the medical education meetings in the 19th century, developing reading, critiquing, analytical, reflective and evaluative skills (Dobrzanska and Cromack, 2005), still hold true as primary, achievable goals of undertaking such an endeavour. The benefits of holding such scholarly activity relate back to the core competencies valued by health professionals, essential in medical research, namely, remaining vigilant of the latest research available in one's field as well as critiquing and appraising the literature in that field, thus bridging research and practice (Kleinpell, 2002, p. 412). Certainly for journal club members, , the opportunity exists to keep aware of the latest research available in a discipline, network with colleagues, and ultimately, by way of promoting these new research findings, improve patient care. For the librarian specifically, it also serves as a way to develop a greater appreciation and understanding of clinical practice guideline development and how pre-established standards are required to ensure that the care provided to cancer patients across the province of Alberta is of consistently high quality.

The success of journal clubs can perhaps be attributed to Brian Haynes, the pioneer of evidence-based medicine. The partnership between the content experts (knowledge management specialists), the clinical experts (nurse facilitators), and the researchers (GURU and the embedded librarian) all play a role in the issues discussed during journal club sessions, ultimately leading to the implementation of guidelines that directly impact patient care. In fact, these roles and responsibilities often intertwine.

Implementation

The idea of launching a journal club with a cancer care team was first envisioned by the Holy Cross librarian in February 2012. From 2009-2010, the librarian had been involved as a preceptor for a journal club with undergraduate medical education students, and wished to apply the concepts learned from this endeavor towards healthcare practitioners. Following a preliminary discussion between the librarian and GURU manager, it was determined that, while all journal club members would take turns gaining the experience of facilitating and leading a session, the librarian would be responsible for taking the lead role and maintaining overall responsibility for this initiative. When the call for participation/establishment of a journal club went out to all research units (~250 staff) working at the Holy Cross site, the librarian had expected there to be greater uptake, although a number of contributing factors, including lack of time commitment and lack of awareness around the functioning of a journal club, may have led to this reluctance in participation.

Dr. Joyce Johnson, registered nurse and director of the Southern California Permanente Medical Group, has developed a pyramid of ten steps to developing a journal club, which was closely adhered to in this pilot study. The fact that journal clubs have existed for over 138 years, speaks to their merit. As numerous other citations discussing journal clubs have done, Johnson (n.d.) posits two fundamental questions for journal club existence, namely "why develop a journal club?" and "are journal clubs really effective?" While a number of Johnson's (n.d.) tips and techniques are cited in several other publications discussing the benefits of holding journal clubs, "there appears to be no well-designed study which has investigated the impact of journal clubs on patient outcomes." This is an interesting argument, considering that several health-related journal clubs undoubtedly deal, whether indirectly or directly, with patient care; although not formally implemented yet, there have been discussions about possibly seeking patient input amongst the provincial Tumour Teams in Alberta, when developing guidelines.

The first two steps of Johnson's (n.d.) pyramid pertain to identifying a leader to organize the journal project and identifying clear goals of the journal club pursuit. The leadership responsibility was held by the Holy Cross embedded librarian, while the goals were identified by the journal club

participants. Johnson's third level focuses on identifying the target audience. In this case, the current journal club is a diversified group, comprised of different roles (i.e. "target audiences").

With respect to scheduling, Johnson's (n.d.) fourth step, the inaugural Journal Club session was held in April 2012 in a virtual library setting referred to as the "Touchdown Suite" on the 6th floor of the Holy Cross Site. Session times were variable (mid-morning or mid-afternoon) and food was provided to encourage collegiality and enjoyment. As Johnson (n.d.) argues, "food is often an important element and supports attendance as well as discussion,". Further, group theory was considered whereby clear boundaries were set, in order to encourage safety and creative thinking (Swift, 2004, p. 68). Reinforcing this facet, each session of the journal club was held at a round table, thus promoting increased participation amongst all members.

The fifth level of Johnson's (n.d.) journal club development pyramid pertains to the type of articles discussed. The article(s) selected (usually two in the case of the Journal Club pilot) aimed to be provocative, so as to stimulate discussion. Because of the variable backgrounds of GURU journal club participants, the opportunity to learn about creation of cancer guidelines from numerous perspectives was facilitated. Newman (2007) states that wherever possible, articles chosen should report original research; while meta-analyses, decision analyses and cost-effectiveness analyses are deemed acceptable, "they are harder to access critically because the results often depend on whether you can trust the authors and their underlying assumptions."

Employing journalistic tendencies, the librarian ensured that the questions of *who, what, where, when*, and *how,* were asked. In other words, *who* was involved in the study, *what* did the study investigate, *where* did it take place, *why* was it conducted, and perhaps, most importantly, how was it conducted? Aztema's (2004) argues that asking these questions are indeed important and despite Journal Club choosing not to adhere to such a formal structure, the librarian ensured that they were considered. Regardless of the informal setting, due diligence on the part of the presenter/facilitator was necessary. When the librarian facilitated and presented a discussion on clinical care pathways, his lack of experience in this area prevented him from forming a biased opinion, a dilemma that can often occur as Aztema admits that it is indeed "easy to criticize" (p. 173). Focusing solely on the facts presented in the article, the librarian maintained objectivity throughout.

In terms of organization, GURU team members kept a record of previously discussed articles in their internal shared network drive. In addition, the librarian created an archive of all articles in JournalFire, an online journal club discussion forum. While this resource has primarily been used as an archival storage facility, due to the demise of this website on October 15, 2013, all citations of article discussed in each Journal Club session have now been transferred over to RefWorks.

Interviews

Upon consultation with the Research Evaluation team at the Holy Cross Site, it was determined that semi-structured interviews would be conducted on an individual basis by a research associate from the Research Evaluation team. To eliminate any potential bias, interviewees were not made aware of any of the questions to be posed prior to the interview taking place. All interviews were held during the summer of 2013. Each session was audio-recorded and lasted approximately 25 to 45 minutes. A summer student was tasked with assisting in the transcription process. All responses were coded and the qualitative data was examined to identify themes and subsequent relationships relevant to the evaluation objectives.

All members of Journal Club, including five staff and the GURU manager, along with the Holy Cross Site embedded librarian (N=7), participated. A semi-structured interview guide was developed to assess participants' overall experiences as well as their comments and reflections regarding their involvement. The 25 questions posed were comprehensive and all-encompassing, addressing several aspects of the journal club process including facilitation, article selection and distribution, individual sessions themselves, involvement in relation to practice, and overall attitudes and perspectives.

Results & Evaluation

Facilitation: Many journal clubs run on a very traditional format, even though discussions may be rather informal. A pre-selected article is introduced, a facilitator/leader provides an executive summary, offers his/her opinion on issues presented, and from there, the discussion can turn into somewhat of a free-for-all, as all are eager to have their say. Johnson (n.d.), in her eighth step, suggests crafting a different format, such as beginning a journal club meeting with a case study, focusing on a particular technique or method found in the pre-required readings about to be presented.

Keeping the channels of communication open was exemplified in the open or round table format discussion in Journal Club. All Journal Club members commented that this approach created an atmosphere where all group members felt comfortable sharing their thoughts. Further, facilitators felt

empowered to encourage open communication with all participants, even on topics that may have been outside one's area of expertise. In addition, there was a high degree of comfort by all participants when asking clarification on unknown concepts, and all were given free reign to openly discuss opinions at each Journal Club session. Nevertheless, it was noted that the level of article discussion varied among certain facilitators, with some adding considerably more input to the conversation than others. Thus, more structured sessions may prove to be beneficial. `

While attendance at each Journal Club was encouraged, adhering to Johnson's (n.d.) step of getting the word out in order to have a meaningful discussion amongst all, it was not mandatory. As a result, competing interests between workload and attendance did take place at times, with pressing deadlines occasionally prevailing over Journal Club meetings.

Three hypotheses pertaining to this pilot were tested to ascertain if perceived needs were met. First that monthly participation in the Journal Club will increase members' knowledge of development, evaluation, and implementation of guidelines. Second, that participants will have acquired a better understanding of the research process and how to critique current guideline research. And finally, that members will report that the Journal Club provided the opportunity to facilitate discussions around topics that are less familiar to them. Following each Journal Club meeting, a group member volunteered to lead the next meeting and thus selected articles to be discussed. This voluntary facilitation allowed for flexibility in team members' schedules, without feeling obligated to assume a facilitative role when workloads were heavy. Due to the diverse backgrounds of the Journal Club members, interviewees indicated that it was challenging to select a topic of interest to the entire group. This led to a series of recommendations: anonymously surveying team members as to specific topics of interest, developing guidelines for use during article discussion, and selecting articles with greater medical science themes. Nevertheless, participants did express notable interest in the diversity of the topics brought forward.

Content: Regarding the themes presented in each session of Journal Club, participants were asked to comment on their level of satisfaction with the topics and subsequent selection of articles for discussion (see Figure 1).

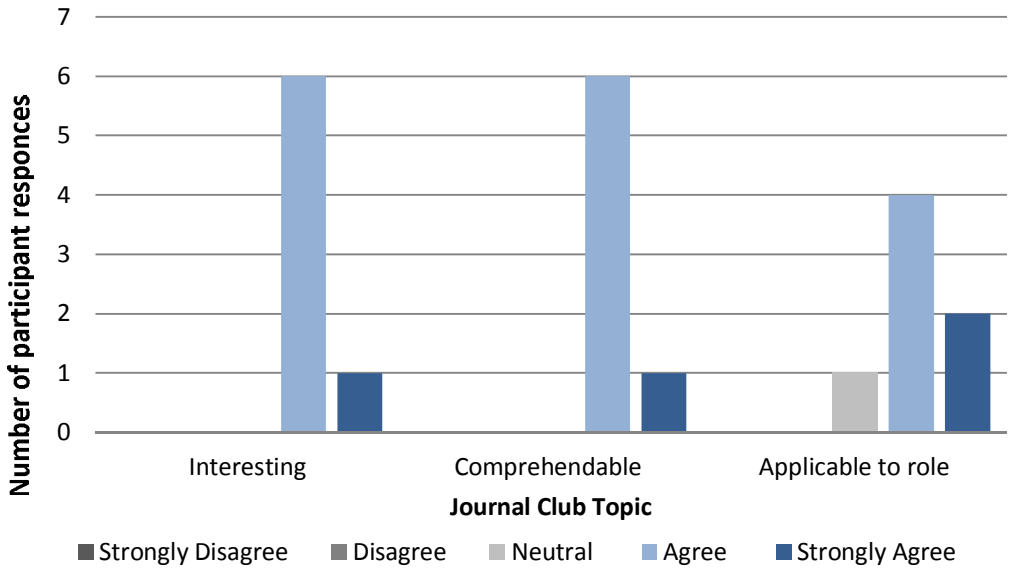


Figure 1. Level of agreement regarding satisfaction with the Journal Club discussion topics.

The overarching purpose of evaluating Journal Club, Johnson's (n.d.) ninth step, was to assess the process, involvement, and impact of this endeavour, in order to determine if the collaborative involvement between knowledge management specialists, nurse facilitators, and an embedded librarian enhanced knowledge sharing, assisted employees in their work, and improved the librarian's understanding of team needs.

Article Selection and Distribution: Choosing an appropriate article for journal club can be a tedious process. A general theme was loosely decided upon before the conclusion of each journal club session, however, there were several aspects that required consideration from article selection to dissemination. Each article was evaluated to ensure it was appropriate for the audience and to determine how the selection of said article(s) would impact a guideline practitioner's work in the field (noting that the librarian is not a content expert). At times, this resulted in a quick search within medical databases and/or journals before a final selection was decided upon. Despite the decision to keep the topic and

subsequent process of identifying article(s) for the next Journal Club rather open-ended, adhering to standard criteria for selecting research articles, as presented in Goodfellow's (2004) paper, is certainly worthy of future consideration. For instance, the desire to determine a specific journal from which articles for discussion are often chosen from, or a pertinent theme, are all decisions that influence the continuation and subsequent success of the journal club. Further, all participants in the Journal Club believe that a committed and dedicated leader made the session worthwhile. Despite these provisions, the literature written on journal clubs does caution that there is no prescribed mandate or guideline for how a journal club should be run.

Buffington, Allen, and Farach (2008)posit a number of factors for determining whether or not a journal club has been successful. These will be examined in relation to the Journal Club. First, a successful journal club should have a minimum two years of existence. At the time of presentation, it will be nearly 20 months since the first session was held. Engagement and desire for the journal club remained high despite two unforeseen circumstances, including a natural flooding disaster that displaced all staff from the Holy Cross Site along with a substantial GURU staffing change. While no sessions were held over the summer months, each participant nonetheless remained active, conducting research, and participating in the evaluation interview. Further, most sessions contained a full contingent of eligible participants, addressing Buffington, Allen, and Farah's (2008) second criteria, namely that more than half of the expected audience should be present at each session.

Often in research endeavors, whether producing an article or reviewing the latest trends in a particular field, examination of the outcomes with respect to the justification of resources and value is undertaken. As Dobrzanska and Cromack (2005) explain, "one way of increasing awareness of current evidence and research findings [sic best practices to improve patient care] is by the introduction of a journal club" (p. 374). When the very idea of launching a journal club was first discussed between the librarian and GURU manager, it was believed that participation in this activity would provide greater understanding in guideline evaluation and development, provide for an opportunity to learn how to critique guideline research, and gain a greater appreciation of the role of grey literature in this venture, particularly as guidelines are, by their nature, a fundamental component of grey literature typology. Further, with one exception, each Journal Club session involved the discussion of no more than two articles. While some facilitators were more lenient than others in allowing certain article discussions to go beyond a pre-determined time allotment, the review of two articles does mirror Aztema's thoughts that a lengthy reading list will undoubtedly lead to "superficial treatment of the studies." (p. 173). Figure 2 describes the participant responses regarding the timeliness of article distribution and the appropriateness of the volume of materials provided.

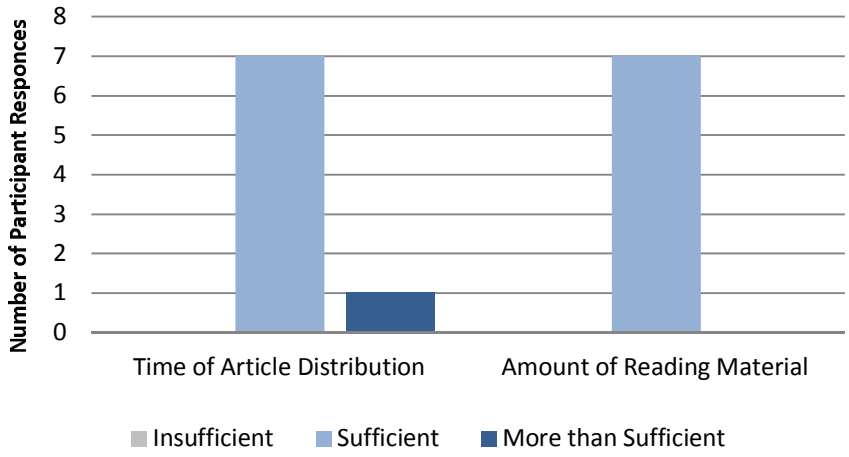


Figure 2. Participant responses regarding timeliness of article distribution and the appropriateness of the quantity of reading material provided.

Impact

When determining the effectiveness of a journal club, Deenadayalan (2008) comments that preparation is essential in ensuring that each meeting runs smoothly and on schedule. When evaluating this trait, it is important to focus on and evaluate lessons learned from the club in lieu of the process of how the club was run (p. 902). Further, as appears to be common practices amongst journal clubs in any discipline, "the initiator of choice of papers was mostly the facilitator…" (p. 903). This aspect may well be the most critical feature of any journal club as article selection and relevance to participants is "a key element of a successful journal club in order to improve reading and critical appraisal skills and knowledge." (p. 905).

It was thought that participation in Journal Club could expand the team's understanding of an embedded librarian's capacity to assist in their work while also providing the librarian an opportunity to engage with teams in a unique format outside of the routine role. For librarians, involvement in team building and learning may enable them to further understand the needs and focus of the team they assist. In addition, team members may become more aware of how an embedded librarian may facilitate their work. This pilot project sought to explore this relationship in an informal Journal Club format.

Assessing the discussions from the past year's Journal Club, the librarian compiled a listing of articles which discussed strategies for conducting successful journal clubs. The launching of the Journal Club, and the subsequent monthly meetings, was certainly not a component of any participant's job description. Further, there was no obligation for a member to attend any or all of the sessions. However, as evidenced by the marked-up pages, the multi-colored highlighting, and the enthusiasm in which the readings were discussed, it was quite clearly evident that all enjoyed partaking in this activity. Devoid of a medical background, the librarian was, in at least a few instances, somewhat overwhelmed by the level of complexity present in some of the articles chosen for discussion. As a result, a second reading "to forage details" (Atzema, 2004, p. 169) was often required, particularly to associate the text with the numerous tables and diagrams that clinical articles often contain. In offering her suggestions on how to read through a paper for journal club, Atzema (2004) presents her perspectives in the IMRAD (introduction, methods, results, AND discussion) format, a method commonly accepted in medical and scientific analyses. While a critical appraisal of the article being discussed was likely never the intent of the Journal Club, critiquing the author's viewpoints and assessing the quality of writing, in much the same way as guidelines are evaluated, often took place.

Johnson's (n.d.) vision of a librarian's role in selecting articles is interesting, yet one that is often true in many journal club situations. According to literature on this topic, it is quite rare that a librarian, particularly one who does not hold a medical background, would be given free-reign, not only to suggest articles/topics of interest for subsequent sessions, but also facilitate and lead meetings. The embedded librarian was very appreciative of being given this opportunity. Journal Club provided an environment for all participants to engage in discussions surrounding subjects both within and outside their area of expertise, as well as learn from fellow colleagues. The opportunity to participate in collaborative, constructive conversation regarding topics allowed team members to share ideas and increase awareness of the knowledge, interests, and opinions of all who participated in Journal Club. This allowed the dispersal of ideas beyond the journal club setting, with several participants commenting that they were able to connect with other GURU team members as well as the embedded librarian to broaden their knowledge base. Thus, even a topic outside of one's practice area still had a positive impact, albeit indirect: "I think it has improved, in just my background knowledge…the discussion often sparks how we apply it to our roles and cancer care in general and I found it really rewarding and truly enhanced my knowledge…it definitely influences my practice" (Watson, 2013).

Newman (2007) offers a few recommendations on identifying bias and/or potential flaws in the study, as a design error can have a trickle-down effect on the feasibility of any guideline produced. With a cautionary tone, Newman concludes by reminding the reader that "the most important part of the discussion is the 'bottom line'. With almost every Journal Club meeting, the discussion was brought full circle, re-emphasizing the key points and generating ideas with respect to applicability to guideline development/implementation . In addition, articles were examined and critically appraised with respect to population size, data analysis, feasibility, general layout, applicability, and so on…

Upon examination of the literature regarding the organization of journal clubs, there does not appear to be a single common guideline detailing the process by which a journal club should run and function (Deenadayalan et al., 2008). In a systematic review of journal club effectiveness reported on by Deenadayalan and colleagues in 2008, "no paper reported on the translation of evidence from journal club into clinical practice" (p. 898). This appears to be a hypothesis that is somewhat open to interpretation, particularly since the entire existence of Journal Club is centered on how the articles discussed in each session better inform guideline developers and tumour group facilitators in interpreting and creating practice guidelines, a very clinically-oriented approach. Further, Deenadayalan (2008) indicates that following an extensive analysis of major medical specialties comprising journal clubs, there was no identifiable journal club devoted to cancer care.

While the participation of a librarian in a journal club is not unique, the level of involvement and engagement with the embedded librarian in Journal Club certainly is. The Holy Cross librarian attended each and every Journal Club since its inception. Members of the GURU team commented that the librarian was already knowledgeable and more than competent at meeting information needs prior to the establishment of Journal Club. Nevertheless, Journal Club allowed for GURU team members to more clearly understand the librarian's role and established greater comfort when asking for assistance.

This involvement increased the librarian's ability to offer enhanced support to the team due to an increased understanding of the team's goals and needs: "I think it helped him with keeping his finger on the pulse of the group. But his influence I think was bigger in terms of just his knowledge of research and literature…" (Watson, 2013).

Despite already hectic schedules, most interviewees were motivated to continue participating in Journal Club. In fact, the opportunity for team building emerged as a key theme throughout the interviews. The ability to network with fellow group members and the embedded librarian, as well as learn about colleagues' background and expertise was seen as a significant benefit of this pilot study. Although educational and research backgrounds varied considerably amongst Journal Club members, this was not perceived as a barrier. In fact, the lack of formality that Journal Club imposed created an environment for open discussion, and all were comfortable discussing topics in an open, non-judgmental manner. While participants did note that a few guidelines surrounding expectations would have been helpful, the relaxed and approachable format of Journal Club enhanced discussion and continued participation well beyond a monthly one-hour timeframe.

Future Considerations

While the Journal Club evaluation period officially ran from April of 2012 until the spring of 2013, meetings are still ongoing. Now that all members have each had an opportunity to facilitate and lead a session, the opportunity exists to expand the journal club concept and invite guest speakers to participate as well. The provincial health structure in Alberta is still relatively new, with Alberta Health Services forming in 2009, and the Knowledge Resource Service this past year. This provides numerous opportunities for Journal Club to partner with new teams, or new journal clubs could take hold across the province, using this pilot study as a guide. Perhaps a virtual approach in the form of a speaker's panel, a successful concept that was tested in the Touchdown Suite in 2011 during Open Access Week is a possibility worth considering.

As with any pilot study of this nature, there are undoubtedly a few limitations that would needed to be addressed for future journal club sessions. Although no formal structure was imposed for each session, thus providing the opportunity for open and free discussion, responses received from participants during the interview process indicated that more explicit directions as to the objectives and purpose of this particular journal club would have been helpful. In addition, the Journal Club may gain even greater validity by inviting an executive director (perhaps as guest facilitator) to a future session; if the director deems this activity to be of high value, it will further entice all participants to continue attending, even if attendance is not mandatory. Further, some members stated that a provision of a list of key questions raised from the readings, distributed a few days before a session would have helped guide and critique the articles (Goodfellow, 2004).

Despite the relaxed and informal atmosphere of Journal Club, a few recommendations were brought forth that may further enhance the awareness of guidelines as a fundamental type of grey literature, as well as suggestions on how to conduct future Journal Club sessions. Clearer expectations, perhaps using a set of guidelines for facilitators would allow for greater article description, pre-determined discussion questions, and personal anecdotes of why a particular article was chosen. Further, the composition of Journal Club varied considerably since its launch, due to staff leaves, workplace reassignments, etc. Thus, brief quarterly surveys may better assist facilitators to tailor topics according to group interests. Finally, collaboration, drawing on each other's strengths, is a fundamental pedagogical principle and an excellent way to use networking to increase awareness of a concept, idea, etc. While there were some barriers in play throughout the Journal Club pilot (article storage, workload, and scheduling), enabling the participation of all interested team members will lead to enhanced team building and improved discussion and learning.

In today's field of librarianship, engagement with clients being served is encouraged and highly recommended. Journal Club provided this opportunity, and the findings from this pilot study support the notion that engaging in the intellectual discussions that Journal Club provided greatly impacted each participant's role. Participating in Journal Club thus helped create a better understanding of the value of grey literature in both research and practice, especially in creating, implementing, and disseminating clinical practice guidelines.

References

Alberta Health Services (AHS). (2013). *Mission & Strategic Direction.* Retrieved November 12, 2013 from http://www.albertahealthservices.ca/190.asp

Australian Research Council (ARC). (2012). *Enhancing the Transparency and Accessibility of Informally Published Research and Information*. Retrieved August 5, 2013 from http://greylitstrategies.info/sites/default/files/docs/Grey_Lit_Strategies_ProjectOverview_August2012.pdf

Aztema, C. (2004). Presenting at journal club: A guide. *Annals of Emergency Medicine, 44*(2), 169-174.

Buffington, D., Allen, J., & Farach, J. (2008.). *Conducting a Journal Club*. Retrieved November 12, 2013 from http://copnt13.cop.ufl.edu/doty/pep/buffingtonffw2008.ppt

Cave, M.T., & Clandinin, D.J. (2007). Revising the journal club. *Medical Teacher, 29*, 365-370.

Deenadayalan, Y., Grimmer-Somers, K., & Kumar, S. (2008). How to run an effective journal club: A systematic review. *Journal of Evaluation in Clinical Practice, 14*, 898-911.

Dobrzanska, L., & Cromack, D. (2005). Developing a journal club in the community setting. *British Journal of Community Nursing, 10*(8), 374-377.

Goodfellow, L.M. (2004). Can a journal club bridge the gap between research and practice/ *Nurse facilitators Educator, 29*(3), 107-110.

Guideline Utilization Resource Unit (GURU) (2013a). *Cancer Guidelines.* Retrieved July 22, 2013 from http://www.albertahealthservices.ca/cancerguidelines.asp

Guideline Utilization Resource Unit (GURU) (2013b). *Handbook*. Retrieved July 22, 2013 from http://www.albertahealthservices.ca/hp/if-hp-cancer-guide-utilization-handbook.pdf

Johnson, J. (n.d.). *Ten Steps to Developing a Successful Journal Club*. Retrieved August 14, 2013 from http://www.freepptdb.com/details-ten-steps-to-developing-a-successful-journal-club-12628.html

Knowledge Resource Service (KRS). (2013). *Cancer Guidelines*. Retrieved August 7, 2013 from http://krs.libguides.com/cancerguidelines

Luby, M., Riley, J.K., & Towne, G. (2006). Nursing research journal clubs: Bridging the gap between practice and research. *Medsurg Nursing, 15*(2), 100-102.

Newman, T. (2007). *Suggestions for Leading a Journal Club*. Retrieved August 20, 2013 from www.epibiostat.ucsf.edu/epidem/personnel/newman_document_repository/JOCLUB_12Jun07.doc

St. Pierre, J. (2005). Changing nursing practice through a nursing journal club. *Medsurg Nursing*, *14*(6), 390-392.

Strain, K.L. (2011). *Embedded Research Librarianship Pilot Project: Evaluation Report.* Alberta Health Services: Public Health Innovation and Decision Support – Evaluation Team.

Swift, G. (2004). How to make journal clubs interesting. *Advances in Psychiatric Treatment, 10*, 67-72.

Watson, L. (2013). *Embedded Librarianship Journal Club Pilot Project: Evaluation Report*. Alberta Health Services: Research and Innovation-Evaluation Team.

GreyGuide - Guide to Good Practice in Grey Literature: A Community Driven Open Resource Project

Stefania Biagioni and Carlo Carlesi, ISTI-CNR, Italy;
Joachim Schöpfel, University of Lille, France;
Dominic Farace and Jerry Frantzen, GreyNet International, Netherlands

Abstract
The goal of this project is to develop an open source repository of good practices in the field of grey literature. That which originated in monographic form will now open and expand to include content from the global grey literature community. Such practices will range from the production and processing of grey literature through to its distribution, uses, and preservation. The repository will contain guidelines such as those in handling theses and dissertations, how to write research reports, metadata required for working papers as well as good practices in the subject areas of agriculture, health, education, energy, environment, et cetera. The purpose of an online repository of good practice in grey literature will provide the many stakeholders in government, academics, business and industry with the benefits of experience, sustained management, and proven results.
The procedure initially applied in this project deals with the design and development of a template that will capture data and information about published as well as proposed good practices within a standard format. While the metadata captured in the template are indeed standardized, their accompanying full-text documents need not be. Furthermore, the template seeks to identify intended users of a good practice, as well as metadata that will facilitate the search and retrieval of records in the repository. Technical developments related to the design and construction of the repository, its eventual platform as well as its maintenance are other related issues addressed in the project. While there are no direct costs associated with the project, each partner is committed to allocate human and material resources needed to carry out their related tasks.
It is expected that the initial phase in acquiring records for the repository will rely on channels available through the Grey Literature Network Service. Populating the repository will be somewhat time-consuming and the first harvest will not produce an abundance of records. The project is long term; however it is all the more worthwhile. The GreyGuide will provide a unique resource in the field of grey literature that is long awaited and which responds to the information needs of a diverse, international grey literature community.

Background
The lead-up to the GreyGuide can perhaps be traced back to the Seventh International Conference on Grey Literature (Nancy, 2005) entitled "Open Access to Grey Resources". At this conference, a proposal was adopted for uniform requirements in the production of grey literature reports[1]. The following year, the Grey Literature International Steering Committee, GLISC[2] was established and GreyNet became one of its first members. Some years later, a monograph was published containing seventeen chapters each mirroring a good practice in the field of grey literature[3]. In that work, one particular chapter dealing with teaching grey literature offered recommendations for best practices in grey literature education[4]. Since then, one can also find on GreyNet's LinkedIn Discussion Group[5] persons who are either in search of a good practice or who are eager to alert the group to a particular practice they consider of value in the field of grey literature. In early 2013 and in response to the Call for Papers for the Fifteenth International Conference on Grey Literature entitled "The Grey Audit"[6], the idea to create a repository of good practices in grey literature was born and the project team was formed. The publisher of the monograph on grey literature was contacted in an effort to interest them in the project. We felt that an online repository would be a better investment in human resources than the publication of a second edition of the monograph. However, the publisher felt that this was not their core business to which we decided to independently construct a repository.

Project Goals
The initial goal of the project was to develop an open source repository of published good practices in the field of grey literature. During the course of the project a number of issues were raised requiring decisions that both expanded the project goals while at the same time further defining the importance that this resource could contribute to the field of grey literature. First of all, the term 'good practice' was opted instead of 'best practice'. The latter term has a connotation, which to potential contributors might set a barrier to the submission of a record. One could read in the term that there is only one best practice related to a particular aspect of grey literature. However, due to the scale and diversity of

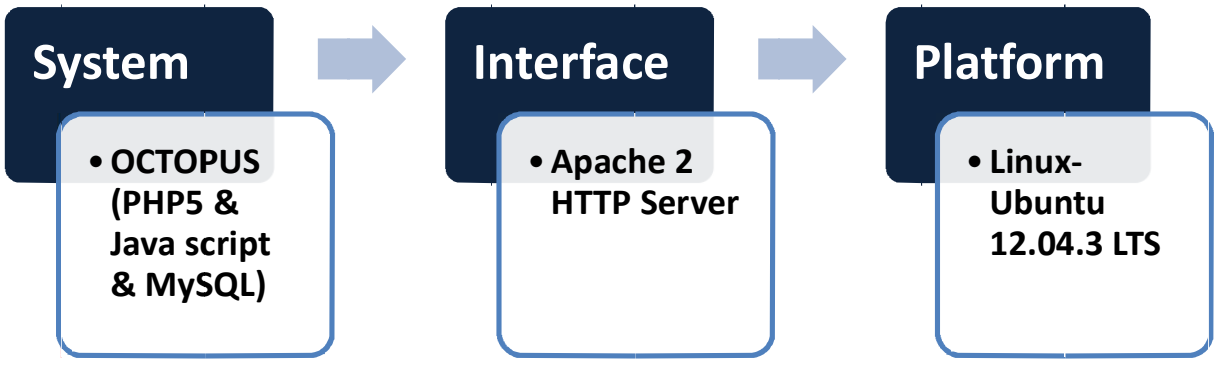
producing bodies, document types, and distribution channels such an interpretation would be limiting. Hence, the term 'good practice' was decided and came to be defined as "simply a process or a methodology that represents the most effective way of achieving a specific objective."[7] The second issue considers that if proposed good practices were accepted alongside published good practices, then the GreyGuide repository would stand to gain more metadata records early on and the proposed good practices would later become published an openly accessible in the repository. And finally, if proposed good practices were openly accessible to wider audiences, then the GreyGuide would function both as an open forum as well as a repository. Potential benefits of this open forum would help to bring researchers and educators as well as other stakeholders in the grey literature community together irrespective of geographical borders.

Method of Approach

The first step was to search the internet for existing templates that could be further adapted and styled for use in the project. Our search produced two templates, which were suited for such purposes. The first template was produced by FAO[8], Food and Agriculture Organization of the United Nations and the second by IEEE[9], Institute of Electrical and Electronics Engineers. Further information was also compiled from the Wikipedia[10]. A template was then drafted and the project technician incorporated it in an online format. The initial test resulted in the division of the one template into two online templates corresponding to published and proposed record entries. The early test indicated that those submitting proposed good practices were confronted with too many record fields that did not apply to their proposal. This could be frustrating to the record creator and a reason not to complete the submission process. The two online templates were then further designed to appear in logical sections corresponding to information about the good practice, the intended users, and the creator/contributor of the good practice. In the case of published good practices – the path, format, and document file name belong to fields in the last section of the template. Each record field was then tagged according to type of response: standard, compulsory, non-compulsory, repeatable, and/or system generated.

Prior to the test phase of the GreyGuide, lead texts were added to each of the repositories webpages providing users with pertinent information on what they would find in that part of the GreyGuide. Embedded lists were also added in order to facilitate record entry. These included a list of grey literature document types borrowed in part from GreyNet's website[11], a list of keywords drawn-up from a number of sources extracted from published titles and indexes available in the conference series on grey literature[12] as well as from The Grey Journal[13], etc. Also included was a list of author's names and affiliations extracted from the Who is in Grey Literature[14]. Two other short lists were further compiled and added: target sectors[15] and target audiences[16]. And finally, help screens were then linked to corresponding record field names providing needed explanations and examples for those creating records.

Technical Notes



Some technical information about the system, interface, and platform are available in the diagram above.

Project Promotion and Community Involvement

During the very formation of the project committee and the formulation of project goals, it was understood that the grey literature community should be informed and involved as much as possible throughout the process. This could be accomplished through promotional activities carried out via existing communication channels - the first of which was the submission of the abstract to the GL15 Call-for-Papers. Once the submission was formally accepted by the conference program committee, a

competition was held in which GreyNet members and non-members were invited to submit a graphic logo that would brand the project and subsequent repository.

GreyGuide

Guide to Good Practice in Grey Literature

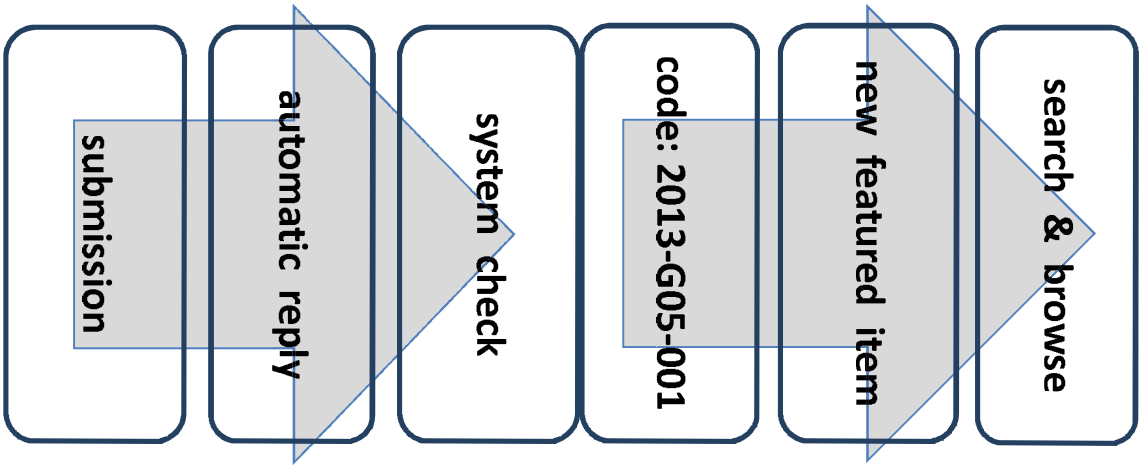
After the project committee selected the current logo from among the three entries, this was then used in the design of a webpage on GreyNet's website bearing the title, GreyGuide Repository [17]. This allowed then for a clear point of reference that could be linked and cross-linked to other social media including GreyNet's Distribution List and Listserv [18], the LinkedIn Discussion Group [19], Twitter [20], etc.
With the above measures in place, a number of postings and news articles also appeared in GreyNet's serial publications: GreyNet Newsletter [21], The Grey Journal [22], GL15 Conference Program Book [23] as well as in other serial publications with interest in the GreyGuide *i.e.* D-Lib Magazine [24], Grey Literature Strategies [25], etc.

Project Meetings and Test Launch

During the months both prior to and after the test launch of the GreyGuide repository, periodic skype calls were made between the project partners. These communications fostered the work and were no less important than the three actual meetings which took place in the spring, summer, and fall of 2 013. The first of such meetings was held in Bratislava (SR) on April 26, 2013 directly following the GL15 Program Committee meeting. After this meeting it was decided to divide the original template into two templates, one for published good practices and the other for proposed good practices. The second meeting was held in Pisa (IT) on July 1st thru 3rd 2013 and focused on the tagging of record fields, the formulation of lead texts, the construction of embedded lists, as well as help screens. There was has te in completing this work due to the scheduled test launch of the repository that would take place at the Library of Congress later that same month. The third and final meeting prior to the actual launch took place in Nancy (FR) on October 11, 2013 and dealt primarily with problems experienced in uploading documents either due to the non-compatibility of some internet browsers or managing character sets, where umlauts and other special characters produced distortions. This meeting was held in conjunction with another meeting on that same day aimed at the establishment of GreyNet's Resource Policy Committee to which the GreyGuide Repository would become integrally involved.

On July 25th 2013, the test launch of the GreyGuide Repository took place during a summer workshop on grey literature[26] held at the Library of Congress in which some 30 federal librarians attended. During the test launch, each participant had access to his/her own computer and were asked to enter either proposed or published good practices that they or their organizations would consider worthwhile for the grey literature community. To assure those participating in the test launch that their records would not be entered in the system should they so choose, they were instructed to simply indicate in the "Note field" - *Not for publication*. During the two hour session, dozens of test records were created and a list of comments and recommendations was drawn-up and sent to the project workers in Pisa. This feedback contributed to the fine-tuning of the GreyGuide repository and was reported back to those federal librarians on November 6, 2013 as part of a live webcast during FEDLINK's annual Fall Exposition[27].

Record Feed and System Flow



On December 2nd 2013, during the official launch of the GreyGuide Repository[28] a demonstration of a record feed into the system and its flow through to completion is shown in the above illustration. In brief, once a record is completed online and submitted to the system the creator of the record receives an automatic reply. The system manager then checks the record and if all of the required fields are completed a unique code is assigned the record and the record then appears for a period of time in the New Feature Items Section. The record is likewise available for browsing, search and retrieval. Should there be a complication with a record; the system manager would notify the record creator before a system code is assigned.

The Way Forward

As with any repository, technical developments are ongoing. Likewise, the acquisition of records requires a proactive policy and relies on the awareness and willingness of the information community it serves to populate the repository. The GreyGuide project partners are intent on the acquisition of good practices in grey literature and will seek to coordinate efforts via GreyNet's newly expanded infrastructure[29] involving three of its committees: the Resource Policy Committee geared to open access, the Community Management Committee focusing on social media, and the LIS Committee dealing with education and training of both students and faculty in the field of grey literature. While the acquisition of records is expected to be time-consuming and the initial harvest will not produce an abundance of records, the project is long term and the benefits worthwhile.

Once a significant number of records are deemed accessible in the GreyGuide, a content and data analysis will be carried out – the results of which should indicate worthwhile uses and applications of this online, open resource and forum for the grey literature community.

References

[1] De Castro, P. and S. Salinetti (2005), Uniform Requirements for grey literature: proposal for the adoption of 'Nancy style', http://hdl.handle.net/10068/697810

[2] http://www.glisc.info/

[3] Grey literature in library and information studies (2010) edited by D.J. Farace and J. Schöpfel http://www.degruyter.com/view/product/41302?rskey=QPXwGs&result=5

[4] Rabina, D. (2010) Course and Learning Objective in the Teaching of Grey Literature: The Role of Library and Information Science Education. http://www.degruyter.com/view/product/41302?rskey=QPXwGs&result=5

[5] http://www.linkedin.com/groups/GreyNet-3718857

[6] http://www.textrelease.com/gl15program.html

[7] Identifying and Sharing Good Practices, SDC Knowledge Management Toolkit, 2004 http://www.fao.org/knowledge/goodpractices/gp-definitionsandcriteria/gp-definitions/en/

[8] http://www.fao.org/knowledge/goodpractices/en/

[9] http://wiki.ieee-earth.org/Template:Submit_Practice

[10] http://en.wikipedia.org/wiki/Best_practice

[11] http://www.greynet.org/greysourceindex/documenttypes.html

[12] http://www.textrelease.com/publications/proceedings.html

[13] http://greynet.org/thegreyjournal/previousissues.html

[14] http://www.textrelease.com/whois2013.html

[15] Academic, Business, Government, Intergovernmental Organization, Non-Governmental Organization, Research Centre, and/or Other.

[16] Instructor, Librarian, Publisher, Author, Researcher, and/or Other.

[17] http://www.greynet.org/greyguiderepository.html

[18] listserv@greynet.org

[19] http://www.linkedin.com/groups/GreyNet-3718857

[20] @GL15conference

[21] http://www.textrelease.com/publications/newsletter.html

[22] http://greynet.org/thegreyjournal.html

[23] http://www.textrelease.com/gl15program.html

[24] http://www.dlib.org/dlib/september13/09inbrief.html

[25] http://greylitstrategies.info/news/greyguide-guide-good-practice-grey-literature-community-driven-open-resource-project

[26] http://www.greynet.org/images/Grey_Boot_Camp.pdf

[27] http://www.loc.gov/flicc/Exposition/2013/Fall/agenda_renFall2013.pdf

[28] http://greyguide.isti.cnr.it/

[29] http://greynet.org/images/GreyNet_2014_Infrastructure_.pdf