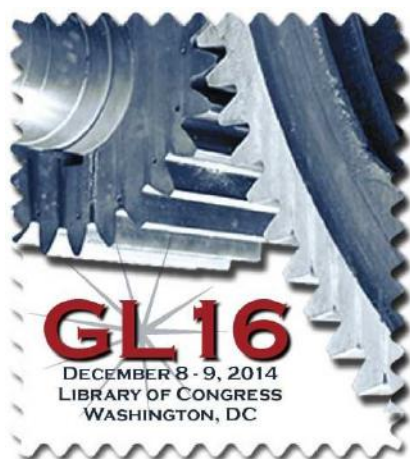# Sixteenth International Conference on Grey Literature

The Library of Congress, Washington D.C. USA, 8-9 December 2014

GL16
December 8 - 9, 2014
Library of Congress
Washington, DC

Program Book

ISSN 1385-2308

Host and Sponsors

LIBRARY OF CONGRESS FEDLINK

CENTRUM VEDECKO-TECHNICKÝCH INFORMÁCIÍ SR

Nuclear Information Section
NIS - IAEA
International Atomic Energy Agency

NEW YORK ACADEMY OF MEDICINE · 1847 · HIPPOCRATES

Inist CNrs

DANS

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

SWINBURNE UNIVERSITY OF TECHNOLOGY

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

KiSTi
www.kisti.re.kr

IIa | Information International Associates

EBSCO PUBLISHING

CIP

Foreword

GREY LITERATURE LOBBY

Engines and Requesters for Change

Decision and policy makers need to be informed on the value and wealth of grey literature, thus legitimizing further investments in this field of information. Lobbying grey literature has its very roots in this international conference series, which has grown and rallies over the past two decades by promoting research and publishing their results. The grey literature lobby seeks to guarantee that the interests of a diverse and widespread community of information professionals and practitioners are served.

Like other fields in library and information science, technology is one of the primary engines driving change in grey literature. However, there are other engines for change that are needed to further sustain and develop this field of information. Policy development and economic stimulus are two such challenges now facing the community.

Changes in the production and sharing of knowledge, changes in the requirements for storage, access, and preservation of grey literature, as well as ensuing change in the demands of users require a concerted effort and response on the part of the grey literature community. New stakeholders and net-users must be summoned and heard, because their requests will further fuel the right kind of debate needed to open the wealth of grey literature to wider audiences.

The social impact of grey literature will be judged by the value placed on its public service. Some thirty-five presentations from authors and researchers from fifteen countries worldwide are harvested in this Program Book.

Dominic Farace   *Amsterdam,*
GREYNET INTERNATIONAL   DECEMBER 2014

GL16 Conference Sponsors

FEDLINK, USA
Federal Library Information Network
Library of Congress

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

GL16 Conference Sponsors

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

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

IIa, USA
Information International Associates, Inc.

DANS, The Netherlands
Data Archiving and Networked Services,
Royal Netherlands Academy of Arts and Sciences

GL16 Conference Moderators and Chair

Moderator Day One

Amanda J. Wilson, Director
National Transportation Library
U.S. Department of Transportation

Conference Chairman

Blane K. Dessy,
Deputy Associate Librarian
Library of Congress

Moderator Day Two

Michele Masias,
Chief Librarian, Justice Libraries
U.S. Department of Justice

Amanda became Director of the National Transportation Library (NTL), an all-digital library, in December 2006. In her time at NTL, Ms. Wilson has focused on building collections, increasing awareness of NTL collections and services, and building a national network of transportation libraries. Under her leadership, NTL has engaged in new initiatives in pursuit of these goals, establishing the Transportation Librarians Roundtable (a monthly webinar series for transportation librarians), providing national coordination of projects between and among regional library networks in transportation, engaging in federal science and technology information leadership through CENDI, and expanding NTL's data programs, including development of a USDOT data catalog and National Transportation Data Archive, and active involvement in the Research Data Alliance and the Federal Statistical Community of Practice and Engagement.
Amanda.Wilson@dot.gov

Blane is a Deputy Associate Librarian with a focus on Planning and Project Management in Library Services at the Library of Congress.
He was the former Executive Director of the Federal Library Information Network at the Library of Congress and was instrumental in establishing the FedGrey Working Group. In 2010, Blane delivered the Keynote Address and at the Twelfth International Conference on Grey Literature. Prior to his appointment at the Library of Congress, Blane had been Director of Libraries at the United States Department of Justice and the first Executive Director of the National Library of Education. Blane came to the Federal Government after working as a State Librarian, Deputy State Librarian, library consultant, and public library director. He was also an adjunct instructor in Management at the Catholic University, School of Library and Information Science.
bdes@loc.gov

Michele serves as the Chief Librarian for the Civil and Criminal Divisions, managing four Justice Libraries. She has over 20 years of library experience in positions that have included serving as the Law Librarian at the Executive Office of the President, the Library Manager at the Defense Technical Information Center, and as a Technical Information Specialist at the Department of Interior Bureau of Reclamation Library. She is currently an elected member of the Fedlink Advisory Board and a Lecturer at the Catholic University of America. She has authored several articles for library publications such as Colorado Libraries, Law Library Lights, Information Outlook, and the U.S. Attorneys' Bulletin. Michele received her MLS from Emporia State University. She's a member of the American Association of Law Libraries and the Law Librarians Society of Washington, D.C.
Michele.Masias@usdoj.gov

Table of Contents

GL16 Program Committee

Blane Dessy **Chair**
Library of Congress
United States

Danielle Aloia
New York Academy of Medicine, NYAM
United States

Ján Turňa
Slovak Centre of Scientific and Technical Information, CVTI SR
Slovak Republic

Stefania Biagioni
Institute of Information Science and Technologies, ISTI-CNR
Italy

Kiseok Choi
Korea Institute of Science and Technology Information, KISTI
Korea

Petra Pejšová
National Library of Technology, NTK
Czech Republic

Dobrica Savic
Nuclear Information Section, NIS-IAEA
Austria

Joachim Schöpfel
University of Lille 3
France

Christiane Stock
Institut de l'Information Scientifique et Technique, INIST-CNRS
France

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

Tomas A. Lipinski
Kent State University
United States

Dominic Farace
Grey Literature Network Service, GreyNet International
Netherlands

GL16 Program Schedule

DAY 1 Monday		DAY 2 Tuesday	
Registration Desk Open *Montpelier Room*	8:30-9:00 AM	*Registration Desk Open* *Montpelier Room*	8:30-9:00 AM
Opening Session *Montpelier Room* Welcome Address, Opening Address, Keynote Address	9:00-10:30 AM	Poster Session and Sponsor Showcase *West Dining Room and Dining Room A*	9:00-11:00 AM
Morning Break	10:30-11:00 AM	*Coffee and Tea Service*	09:00-11:00 A
Session One *Montpelier Room* **"Public Awareness of Grey Literature"**	11:00-12:30 PM	Session Three *Montpelier Room* **"Open Access to Research Data"**	11:00-12:30 PM
Lunch (Set-up Posters) *West Dining Room and Dining Room A*	12:30-1:30 PM	*Lunch*	12:30-1:30 PM
Session Two *Montpelier Room* **"Publishing and Licensing Grey Literature"**	1:30-3:00 PM	Session Four *Montpelier Room* **"Managing Change in Grey Literature"**	1:30-3:00 PM
Afternoon Break	3:00 -3:30 PM	*Break*	3:00-3:15 PM
Introduction to Posters *Montpelier Room* Briefings by Poster Presenters	3:30-4:30 PM	Closing Session *Montpelier Room* Report by Conference Chair and Moderators, Conference Handoff, and Farewell	3:15-3:45 PM
Recommended Tour	4:30 PM	*Recommended Tour*	3:45 PM

DAY ONE

Day Moderator: Amanda Wilson, U.S. Department of Transportation, United States

		9:00 – 10:30 AM
		MONTPELIER ROOM

OPENING SESSION

Welcome Address	*Blane Dessy*, *Library of Congress, United States*	
Opening Address	**The Value of Grey Literature : Finding the Silver,** *Roberta I. Shaffer, United States*	13
Keynote Address	**Where is the evidence?** **Enhancing the role and value of grey literature for public policy and practice** *Amanda Lawrence, Swinburne University of Technology, Australia*	15

10.30 -11:00 AM   Morning Break

	11:00 – 12:30 PM
	MONTPELIER ROOM

SESSION ONE – PUBLIC AWARENESS OF GREY LITERATURE

The Intersect of Cultural Studies with other focus areas in the HDIAC Database   **17**
June Crowe and Crystal Sherline, Information International Associates Inc., United States

Mapping Italian Grey Communities: What Is There Beyond The Academy?   **21**
Silvia Giannini and Stefania Biagioni, Istituto di Scienza e Tecnologie dell'Informazione;
Sara Goggi and Gabriella Pardelli, Istituto di Linguistica Computazionale; CNR, Italy

Shades of Grey: Integrating Metadata for Discovery in a Mixed-Content Single-Subject Library   **27**
Genevieve Podleski, Federal Reserve Bank of St. Louis, United States

How Grey Literature Informs Policy and Decision Making: The Necessity to Understand the Processes   **32**
Bertrum H. MacDonald, James D. Ross, Suzette S. Soomai, and Peter G. Wells, Dalhousie University, Canada

12:30-1:30 PM   Lunch (Set-up Posters)   *WEST DINING ROOM AND DINING ROOM A*

	1:30 – 3:00 PM
	MONTPELIER ROOM

SESSION TWO – PUBLISHING AND LICENSING GREY LITERATURE

Is the Licensing of Grey Literature Using the Full Palette of "Contractual" Colors?   **35**
Tomas A. Lipinski, School of Information Studies, University of Wisconsin—Milwaukee, United States
Andrea J. Copeland, School of Informatics and Computing Indiana University, United States

Publishing Geodesy, Topography and Cartography Research via Invenio   **41**
Jiří Drozda and Veronika Synková, Research Institute of Geodesy, Topography and Cartography;
Petra Pejšová, National Library of Technology, Czech Republic

Free Licences and Creative Commons: A Powerful Tool for Open Access Publishing in Grey Literature   **46**
Petra Pejšová, National Library of Technology, Czech Republic
Marcus Vaska, Knowledge Resource Service, University of Calgary, Canada

Degrees of Openness: Grey Literature in Institutional Repositories   **52**
Joachim Schöpfel, Charles de Gaulle University Lille 3;
Hélène Prost, CNRS, Associate member GERiiCO, France

3:00-3:30 PM   Afternoon Break

	3:30 – 4:30 PM
	MONTPELIER ROOM

INTRODUCTIONS TO POSTERS

On the afternoon of Day One, each person presenting a poster will have the opportunity to introduce their work in the Main Forum. The Poster Session will continue in adjoining conference rooms on the morning of Day Two, where the presenters will be able to meet with conference delegates and participants in an informal setting.
 (See Listing of Posters, page 12)

	4:30 PM

DAY TWO

Day Moderator: Michele Masias, U.S. Department of Justice, United States

9:00 – 11:00 AM	
WEST DINING ROOM AND DINING ROOM A	

POSTER SESSION AND SPONSOR SHOWCASE

Conference posters are scheduled to be presented in an informal setting   **57**

11:00 – 12:30 PM	
MONTPELIER ROOM	

SESSION THREE – OPEN ACCESS TO RESEARCH DATA

Data: Is it Grey, Maligned or Malignant?   **101**
Julia Gelfand and Daniel Tsang, University of California, Irvine, United States

Open data, grey literature and disciplinary differences – Perspectives from a Dutch data archive   **102**
Marnix van Berchum, DANS, Data Archiving and Networked Services, Netherlands

Introduction of "Data Flagging" to utilize numerical data on the Internet - case of JAEA Fukushima Nuclear Accident Archive   **107**
Kiyoshi Ikeda, Mayuki Gonda, Shun Nagaya, Misa Hayakawa, Yukinobu Mineo, Katsuhiko Kuni, Minoru Yonezawa, Keizo Itabashi, Japan Atomic Energy Agency, Japan

Moving towards Open Access to research data in health sciences: a changing landscape for Grey Literature   **113**
Paola De Castro, Istituto Superiore di Sanità, Italy

12:30-1:30 PM   Lunch

1:30 – 3:00 PM	
MONTPELIER ROOM	

SESSION FOUR – MANAGING CHANGE IN GREY LITERATURE

Changes in the International Nuclear Information System   **115**
Dobrica Savić, Nuclear Information Section, NIS; Department of Nuclear Energy, International Atomic Energy Agency, IAEA, Austria

GreyGuide, GreyNet's web access portal and lobby for Grey Literature   **117**
Dominic Farace and Jerry Frantzen, GreyNet International, Netherlands
Stefania Biagioni, Carlo Carlesi, and Roberto Ponti, ISTI-CNR, National Research Council, Italy
Christiane Stock, Inist-CNRS, National Centre of Scientific Research, France

Grey Lit Tutorials   **122**
Rachel Smiley, Sarah Davis, and Debbie Rabina; Pratt Institute, School of Information and Library Science, United States

Publishing and Communication: The GreyNet Community Management Committee   **123**
Joachim Schöpfel, Charles de Gaulle University Lille 3, France

3:00-3:15 PM   Break

3:15 – 3:45 PM	
MONTPELIER ROOM	

CLOSING SESSION

Report by Conference Chair and Moderators,
Conference Handoff, and Farewell

3:45 PM	

GROUP TOUR

To be announced

POSTER SESSION AND SPONSOR SHOWCASE

Analysis of Collection and management of the Korea National R&D Report   **57**

Kiseok Choi, Cheol-Joo Chae, Yong-hee Yae, and Yong Ju Shin,
Korea Institute of Science and Technology Information, KISTI, Korea

An Investigation into the use of Grey Literature in Evidence-Based Medicine: **62**
The role of the NHS Library and Information Services (LIS) Professional

Joanna Hooper, University Hospitals Bristol NHS Foundation Trust, BRI, United Kingdom

55 Years of the National Union Catalog of Manuscript Collections, NUCMC : A Program of the Library of Congress   **63**

Beth Davis-Brown, Peter Goodman, Roberto Sicre, Library of Congress;
Cooperative and Instructional Programs Division, NUCMC Program, United States

An Attempt to Nuance the Understanding of Professional Reports in Archaeology   **64**

Lisa Börjesson, Department of Archives, Libraries, Museums; Uppsala University, Sweden

Publishing and Communication: GreyNet Community Management Committee   **69**

Joachim Schöpfel Chair (France), Cees de Blaaij and Dominic Farace (Netherlands), Amanda Lawrence (Australia),
Hélène Prost (France), and Marcus Vaska (Canada)

ETD4OA, Electronic Theses and Dissertation for Open Access   **70**

Joachim Schöpfel, Helene Prost, Marjorie Piotrowski, University of Lille, France;
Eberhard R. Hilf, Thomas Severiens, Paul Grabbe, Institute for Scientific Networking; University Oldenburg, Germany

Open Access Korea, Phase 1: Five Years On   **73**

Hyekyong Hwang, Seonheui Choi, Hyesun Kim; Korea Institute of Science and Technology Information, KISTI, Korea

Big data availability and full exploitation: **79**
A 'wide screen' overview on potential opportunities and current criticalities

Luisa De Biagi, Massimiliano Saccone, and Roberto Puccinelli, National Research Council, CNR Central Library, Italy

Think Tanks, Twitter, and Grey Literature   **86**

Danielle Aloia, New York Academy of Medicine, NYAM, United States

Capturing from the start: **89**
Managing grey literature in a brand new university and enriching the institutional repository

Mohamed Ba-Essa and J.K. Vijayakumar, King Abdullah University of Science and Technology, KAUST, Saudi Arabia

Marine Planning and Service Platform (MAPS);   **93**
An Advanced Research Engine for Grey Literature in Marine Science

Sara Goggi, Monica Monachini, Francesca Frontini, Roberto Bartolini, Gabriella Pardelli, ILC-CNR, Italy;
Maurizio De Mattei, Franco Bustaffa, DP2000, La Spezia, Italy   and   Giuseppe Manzella, ETTsolutions, Genova, Italy

Using the PressForward Plugin to Discover and Manage Grey Literature   **95**

Stephanie Westcott, Lisa Rhody, Joan Fragaszy Troyano, and Kim Nguyen, The Roy Rosenzweig Center for History and New Media; George Mason University, United States

GreyGuide, Guide to Good Practices and Resources in Grey Literature   **96**

Stefania Biagioni, Carlo Carlesi, Roberto Ponti, ISTI-CNR, Italy

Driving policy development for grey literature – The Pisa Declaration   **98**

Herbert Gruttemeier, Chair GreyNet Resource Policy Committee, Inist-CNRS, France

Providing access to research data on Africa   **99**

Laverne Page, African and Middle Eastern Division; Library of Congress, United States

The recent Italian regulations about the open-access availability of publicly funded research publications,   **100**
and the documentation landscape in astrophysics

Monica Marra, Istituto Nazionale di Astrofisica, INAF; Osservatorio Astronomico di Bologna, Italy

Opening Address

The Value of Grey Literature: Finding the Silver

Roberta I. Shaffer

Roberta I. Shaffer was appointed Associate Librarian for Library Services on January 2012 and she retired from this position in August 2014. As Associate Librarian, she managed 53 divisions and offices whose over 2,000 employees are responsible for acquisitions, cataloging, public service, and preservation activities, services to the blind and physically handicapped, and network and bibliographic standards for America's national library.

Shaffer graduated cum laude from Vassar College with an A.B. degree in political science/demography, with highest honors from Emory University with a master's degree in law librarianship and cum laude with a J.D. from Tulane University School of Law.

Shaffer has a rich and diverse background in academics, management, law, librarianship, and technology. Over the course of her nearly 40-year career, Roberta has worked at University of Houston Law Center, at the Law Library of Congress, as Fulbright Senior Research Scholar, and George Washington University National Law Center Library.
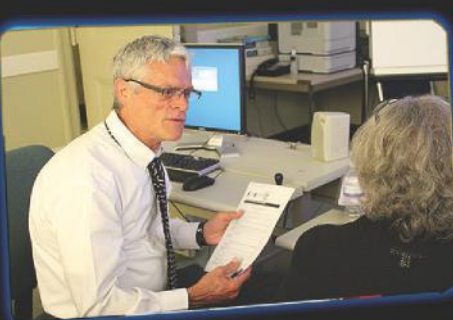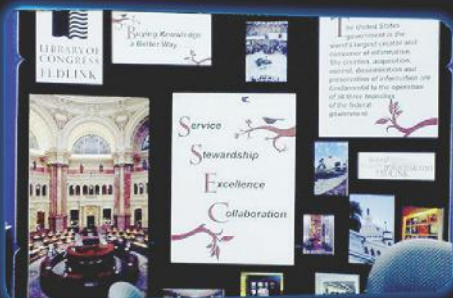
She served concurrently as adjunct coordinator of the law librarianship specialization at the School of Library and Information Science at the Catholic University of America and as director of research services at the law firm of Covington and Burling in Washington, D.C. In 1999, she was appointed and served for two years as dean and professor at the Graduate School of Library and Information Science at the University of Texas at Austin.

In 2002, she accepted the position of director of external relations and program development and professor at the College of Information Studies at the University of Maryland in College Park, where she implemented and ran the Master of Information Management (MIM) degree.

After a 25 year hiatus, Shaffer returned to the Library of Congress in 2005 as executive director of the Federal Library and Information Center Committee/Federal Library Network (FLICC/FEDLINK). In 2009, she was named as the 22nd Law Librarian of Congress.

FEDLINK

an organization
of federal agencies
working together
to achieve optimum use of
resources and facilities
of federal libraries
and information centers
by promoting
common services,
coordinating and sharing
available resources, and
providing continuing
professional education.

LIBRARY OF CONGRESS FEDLINK

101 Independence Ave, SE ~ Washington, DC 20540-4935
FEDLINK Main Number (202) 707-4800
FEDLINK Hotline (202) 707-4900

Keynote Address

Where is the evidence?
Enhancing the role and value of
grey literature for public policy and practice

Amanda Lawrence

Swinburne University of Technology, Australia

Digital technologies have radically increased our capacity to produce and disseminate research and information and many organisations including government departments and agencies, academic centres, NGOs, lobby groups and companies are now engaged in the production and use of digital grey literature. The result has been an explosion in the publication of reports, discussion papers, working papers and many other kinds of grey literature on the internet, transforming the way in which the community is able to access research and information. Yet the rapid pace of change has meant that there is little understanding of the role and value of this material and we still lack appropriate infrastructure, policies and procedures, copyright legislation and resources to adequately manage and provide ongoing access to these valuable resources.

In this presentation I will explore the key role played by grey literature in public policy research and practice based on the results of a 3-year research project supported by the Australian Research Council. As part of this project we have undertaken large-scale online surveys and interviews with users, producers and collectors of research and information, reviewed current policies and infrastructure and considered the international context in which grey literature operates. The evidence supports the conclusion that grey literature plays a vital role in the policy and research environment, equal or greater than that of journals and books, but that its production, collection and preservation is inadequate, resulting in poor productivity, the loss of costly research and duplication of projects amongst other things. The paper will present recommendations for policy and infrastructure changes that could improve the way digital grey literature is collected and made accessible and seek feedback from the audience for the final report due out in 2015. I hope that through this presentation and the research project we can find clearer ways to work together on local, national and international initiatives to enhance the value of grey literature for public policy and practice.

This project is being undertaken by Swinburne University of Technology and Victoria University in partnership with the National Library of Australia, the National and State Libraries Australasia (NSLA), the Australian Council for Educational Research and the Eidos Institute.

Bionote

Amanda Lawrence is a researcher with Swinburne University's Institute for Social Research. Ms Lawrence is managing editor of Australian Policy Online and a guest lecturer in information architecture at the Royal Melbourne Institute of Technology (RMIT). She holds a Graduate Diploma in Library and Information management (RMIT) and BA (Hons) Arts (Melbourne). She is qualified for membership of the Australian Library and Information Association (ALIA).
Email: alawrence@swin.edu.au

The Intersect of Cultural Studies with other focus areas in the HDIAC Database

June Crowe and Crystal Sherline

Information International Associates, Inc. United States

Our bibliometric research will examine a select set of documents in the Homeland Defense and Security Information Analysis Center's (HDIAC) collection. The focus will be on documents in the "Cultural Studies" focus area. There are over 210,000 documents in the HDIAC collection, much of it being grey literature. Staff uses a template that includes bibliographic information, keywords and task areas, as well as other descriptive information to catalog items for inclusion in HDIAC's database. In staff discussions, it was discovered that there are focus areas that span into other focus areas. For example cultural studies overlap into medical, alternative energy, and HDS areas. The purpose of this research is to examine the transects that Cultural Studies has with other focus areas, by examining the keywords and task area terms that include the overarching term "Cultural Studies". This process will assist in clarifying the terms used to identify cultural studies and facilitate the tagging and acquisition process. It will also give staff a model upon which to gain a deeper understanding of how culture intersects with other disciplines. A bibliometric study will be conducted to evaluate the HDIAC database, START, quantifying the Activity Index (AI) of the HDIAC database. Where:

$$AI_1 = \frac{Total\ Number\ of\ Publications\ in\ START}{Total\ Cultural\ Studies\ Publications\ in\ START}$$

And

$$AI_2 = \frac{Total\ Number\ of\ Publications\ in\ START}{Total\ Cultural\ Studies\ Cross\ Publications\ in\ START}$$

This process of bibliometric research will allow HDIAC staff to quantify the multidisciplinary nature of Cultural Studies and provide a vantage point for a qualitative study that examines the intersect with the other focus areas in HDIAC.

Keywords: Scientific data, Information policies, Grey Literature, Data Discovery, Bibliometric Analysis.

Bionotes

June Crowe is the Senior Researcher and Group Manager, Open Source Research Division at Information International Associates, Inc. (IIa). She received her AMLS from the University of Michigan, Ann Arbor, and her M.Ed. in geographic education from the University of Georgia, Athens. Her current position, subject matter expert cultural studies, is with the Homeland Defense and Security Information Analysis Center (HDIAC) in Oak Ridge, TN. Her primary interests are Grey Literature research, socio-cultural and S&T research, digital repositories, and open source intelligence tools. Email: jcrowe@iiaweb.com

Crystal Sherline is a Technical Information Analyst for Homeland Defense and Security Information Analysis Center (HDIAC), a researcher at Information International Associates (IIA), and doctoral candidate at the University of Tennessee, Knoxville. She received a Master's degree from the University of Maryland, and an MLIS from Florida State University, and will be receiving her Ph.D. from the University of Tennessee in July 2014. She has experience in operations of both public and academic libraries. Her interests are varied, but consist of primarily research, including Network Analysis, Grey Literature, HCI, Social Science, and data sharing. Email: csherline@iiaweb.com

The Intersect of Cultural Studies with Other Focus Areas in the HDIAC Database

June Crowe &
Dr. Crystal Sherline

GreyLit Conference 16
December 7-9, 2014

IIa Proprietary Information

INTERSECT OF CULTURAL STUDIES

- HDIAC Overview
- Research Method
- Cultural Studies Intersects
- Taxonomy Development
- Conclusion
- Questions

IIa Proprietary Information

HDIAC OVERVIEW

- S&T Database with over 210,000 items
- Knowledge Center of Excellence for DoD and Homeland Defense and Security
- Formally CBRNIAC
- Now, as HDIAC, has eight focus areas: HDS, CIP, WMD, CBRN, Biometrics, Medical, Cultural Studies, and Alternative Energy

IIa Proprietary Information

RESEARCH METHOD

- Algorithm for literature review in START database
- Research results for Cultural Studies Intersects

Iia Proprietary Information

TAXONOMY DEVELOPMENT

- HDIAC START System
- Need for a more detailed system
- No Existing Taxonomy that fits
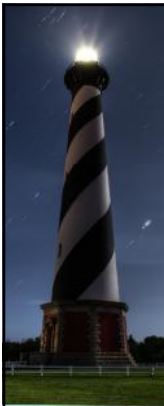- Where we are in the process

Iia Proprietary Information

Contact Us

Information International Associates, Inc.
104 Union Valley Road
PO Box 4219
Oak Ridge, TN 37831-4219

Main (865) 481-0388
Fax (865) 481-0390

www.iiaweb.com

Bonnie C. Carroll
President
bcarroll@iiaweb.com
(865) 298-1220

Kelly Callison
COO
kcallison@iiaweb.com
(865) 298-1268

Iia Proprietary Information

Slovak Centre of Scientific and Technical Information
SCSTI

Achieve your goals with us

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

Mapping Italian Grey Communities:
What Is There Beyond The Academy?

Silvia Giannini and Stefania Biagioni,
Istituto di Scienza e Tecnologie dell'Informazione; CNR, Italy

Sara Goggi and Gabriella Pardelli,
Istituto di Linguistica Computazionale; CNR, Italy

The following title was published on an influential Italian newspaper, *La Stampa*, on November 7, 2013. "Tra i tesori della 'Letteratura Grigia' un'Eneide in napoletano del '600". The article is about the presentation of the "Fondo De Mauro" on the Italian Network of Popular Culture: this fund originates from a private collection built up in several decades and donated to the Network by Tullio De Mauro[1] and his wife Silvana Ferreri in 2011; it is made up of thousands of books, brochures, pamphlets relating to Italian dialects and minority languages.

In the field of healthcare, while browsing on the web, we found the title of a PhD thesis of last year: "Letteratura Grigia nelle meta-analisi delle prove ripartite con scelta casuale degli interventi di sanità".

In the legal environment, the title of a Seminar at the University of Siena emerges from the web: "La "centralità" della legge e la letteratura grigia. Profili di politica del diritto in Italia tra Otto e Novecento"[2].

From these findings, the idea of a survey on the wide variety of grey material available on Italian web portals arises. A first analysis shows that this material is available in different forms and dissemination is carried out through various means such as thematic bibliographies, newspapers articles, various types of documents published in pdf format or simple descriptions on web sites. The following are a few examples excerpted from some home pages belonging to our corpus:

- ≈   Collection of *grey literature*. The Historical Archive of Women candidates for becoming the repository where the memories about these themes will be stored…[3]
- ≈   *Grey literature …* Master copies · Reprints · Unreleased copies · *Grey Literature*. Archives for the history of Education…[4]
- ≈   *Grey literature*. The high quality brand of parks…[5]
- ≈   *…* International *grey literature*; national and international legislative data on the topic of drug addiction and related themes; documentary archive …[6]
- ≈   Besides literature in German, there is literature in other languages and *grey literature* as well – in particular catalogues of museums and exhibitions…[7]
- ≈   It is about a few thousand of books, brochures and documents of *grey literature* concerning two topics, Italian dialects and minority languages…[8]

---

[1] Famous linguist and former Italian Ministry of Education.

[2] The Seminar presents the results of a 2007 PRIN research project funded by the Italian Minister of Research: «Perpetue appendici e codicilli alle leggi italiane». Le circolari ministeriali, il potere regolamentare e la politica del diritto in Italia tra Otto e Novecento.

[3] http://www.archiviodonne.bz.it/it/progetti/raccolta-di-letteratura-grigia?

[4] http://www.historied.net/portal/index.php?option=com_content...id...?

[5] http://ambiente.regione.emilia-romagna.it/parchi-natura2000/consultazione/pubblicazioni/letteratura-grigia

[6] http://www.ceisroma.it/upgrade/biblioteca-agora/

[7] http://artlibraries.net/allg_infos_it/KUNST_GDK.php

[8] http://www.reteitalianaculturapopolare.org/archivio/fondi-in-rete/fondo-tullio-de-mauro.html

Given this scenario, the research aims at verifying whether – and eventually how much – the grey literature available on the web is actually structured, accessible or even managed by systems dealing with its organization and aiming at its retrieval and storing. The utmost goal is to build up a map of non-academic communities and their mechanisms for managing, presenting and disseminating this type of material: a sort of journey among the streams of the web which channel meeting minutes, invites, manifests, fliers, pictures, newspapers articles, journalistic services and audio/video material on various topics. These "grey" products – by conveying basic information about social and popular culture – store, represent and spread knowledge. Significant examples could be identified in the web sites presenting the following matters: history of women's culture and of their movement in some specific Italian regions; projects dealing with the sustainability of urban environment with respect to childhood and adolescence; parks and other natural protected environments; archaeological documentation such as draft reports, diaries from the site, letters and miscellaneous documentation; nursing and health-related disciplines which produce guidelines, diagnostic and therapeutic courses, informative material for patients and their families. Also the theatrical culture is nourished by "grey products" as video archives, collections of music LPs and CDs, brochures, scripts, autograph manuscripts (i.e.. letters, correspondences, fliers, musical scores). In substance, a heterogeneous set of material which could reveal especially interesting to both researchers, scientists, professionals and simple fans and lovers of the various subjects if ever made available and usable.

Focus

Analysis of the documentation and production of taxonomies finalized at the creation of a map of non-academic communities and stakeholders involved in the management of grey material.

Material and methods

The survey examines the several disciplines, the typology, the institutional nature and the fields to which these grey communities belong, noticing as well the variety of documentation provided, the structure of information and the presentation and access modalities through the following steps:

1) Selection of the web portals as resulted from the query "letteratura grigia" OR "letteratura non convenzionale" OR "documentazione grigia" OR "materiale grigio".

2) Creation of a corpus made up of 28,000 occurrences.

3) Analysis of the various communities and of the grey material retrieved.

4) Statistical elaboration of the data.

Conclusions: reflections on the communities and stakeholders involved in the management of grey literature and on the various ways of presenting the documentation provided by the web sites selected for this survey.

Bionotes

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

Stefania Biagioni see bionote page 117

Sara Goggi see bionote page 94

Gabriella Pardelli see bionote page 94

**MAPPING ITALIAN GREY COMMUNITIES.
WHAT IS THERE BEYOND THE
ACADEMY?**

Silvia Giannini, Stefania Biagioni, CNR-ISTI, Pisa Italy
Sara Goggi, Gabriella Pardelli, CNR-ILC, Pisa Italy

Summary

The survey examines Italian Grey Communities beyond the Academy

- Objective
  - Creation of a map of non-academic communities and stakeholders involved in the management of grey material
- Method
  - Focused Google queries and acquisition of results
  - Building of the Corpus
    - Analysis of documentation and categorization (whom, what, how…)
  - Analysis of results and production of taxonomies
- Conclusions and Remarks

The method: search and retrieve

Focused Google queries and acquisition of results

- Selection of the web portals as resulted from the query:
  "letteratura grigia" OR "letteratura non convenzionale" OR "documentazione grigia" OR "materiale grigio"

Google results
  a. Number of indexed documents considered as "interesting" with respect to the terms of the query (about 42000)
  b. List ordered by relevance of URL (never more than 1000)
  c. 800 URLs retrieved for building the sample (Corpus)

Google

GL16 Conference, December 8-9, 2014 Library of Congress, Washington D.C., USA

On the scene

Several different communities involved in the management of grey material, each with quite different
- Institutional nature
  - public and private, ecclesiastic ….
- Fields to which these grey communities belong
  - cultural & social/welfare interest, healthcare, research, education and beyond

Several types of "grey material", each with quite different
- Formats and dissemination
  - various types of documents published in pdf format or simple descriptions on web sites, thematic bibliographies, newspaper articles, multimedia, catalogues, repositories
- Organization, representation, storing, preservation …

GL16 Conference, December 8-9, 2014 Library of Congress, Washington D.C., USA

The method: categorization

The Corpus is built by grouping the gathered data in 12 informative classes assigned to each URL

Classes	Example
Stakeholder	Association
Type of Stakeholder	Social Centre
Name of Stakeholder	"Lilith - Archivi delle donne e del femminismo"
Infrastructure/service	Archive
Field	Social
Topic	Feminism
Type of material	Leaflets, bibliographies, private correspondence, manuscripts, minutes …
Format	html
Access	free
Domain	it
Description	GL appears in the description of the archive property

---

GL16 Conference, December 8-9, 2014 Library of Congress, Washington D.C., USA

Analysis of the data

How do we behave when facing a classified list of results?
- The Corpus has been "cleaned up" based on the relevance of content:
  - exclusion of academic web sites, web sites advertising/selling products of grey colour, urls duplicates…
- The data have been ordered on the bases of the stakeholders' nature:
  - e.g. public, private, business, personal …
- The communities have been collected on the bases of similar features and content:
  - e.g. childhood, music cinema and theatre, women, welfare, professional associations, healthcare …

Which criteria should be used for assembling the communities?
- subjects of relevant social interest
- richness of content
- organisation, representation, storing, preservation of data/documents
- …

GL16 Conference, December 8-9, 2014 Library of Congress , Washington D.C., USA

Stakeholders

Who is most interested in GL in Italy beyond the Academy?

There are stakeholders of public, private and hybrid nature

- The public sector discloses information homogeneously because it should respond to institutional standards as far as the treatment of data is concerned:
  - State entities and agencies
  - Public Institutions/Ministries/Regions/Municipalities
  - Clerical entities
  - Research centres
  - International agencies

- The private sector discloses information in a heterogeneous way according to its own purposes:
  - Companies/corporations
  - Public and personal blogs
  - Personal homepages
  - Associations/organizations/foundations

- The hybrid sector (public/private)
  - Associations/foundations
  - Spin-offs

GL16 Conference, December 8-9, 2014 Library of Congress, Washington D.C., USA

Communities

Significant examples have been identified on the web sites presenting a heterogeneous set of material which could prove to be especially interesting to researchers, scientists, professionals and simple fans and lovers of the various subjects if ever made available and usable.

| Environment | Women | Health | History |

Several communities handling and disclosing "grey" documentation in various ways come to light from this analysis. The ensemble of information collected allowed to identify and delimit 4 illustrative communities.
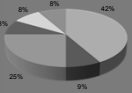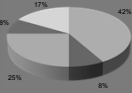
GL16 Conference, December 8-9, 2014 Library of Congress, Washington D.C., USA

Communities - Stakeholders

ENVIRONMENT
- **Sustainable cities**
- **Geomorphology**
- **Mediterranean sea**
- **Aero-spatial research**
- **Mussels' production chain**
- **....**

8% 8% 42%
8%
25% 9%

Association/Organization
Company/Society
Public Body
Region
Ministry
International Organization

WOMEN
- **Archives of the history of women**
- **Voices, visions and actions of women**
- **Violence on women**
- **Archives of feminism**
- **....**

17% 42%
8%
25% 8%

Association/Organization
Ministry
Municipality
Private
Region

GL16 Conference, December 8-9, 2014 Library of Congress, Washington D.C., USA

Communities - Stakeholders

HEALTHCARE
- **Drug dependences, mental diseases and youth disorders**
- **Work, safety and social politics**
- **Alcohol-related pathologies**
- **Eating disorders**
- **....**

Association/Organization · Company/Society · Ministry · Municipality · Public Research · Region

HISTORY
- **History of the Church**
- **History of the Italian Risorgimento**
- **Archives for the history of Education**
- **Historical archives of political parties**
- **....**

Association/Organization · Company/Society · Municipality · Public Research · Region

---

GL16 Conference, December 8-9, 2014 Library of Congress , Washington D.C., USA

Conclusions & Remarks

Using Google-search as a source for building the GL Corpus has been useful for evaluating the state-of-the-art …

- Disadvantages
  - There is an "explosion" of information … who is producing it? a Person, a Society, a Region, a Foundation?
  - Pages containing large amounts of unrelated material (downloads, graphics, non-pertinent material …)
  - Many web sites cannot be accessed / change their URLs / are not updated / information is either cancelled or duplicated
  - It is impossible to have more than 1000 results

- Advantages
  - Google finds the unobtainable…
  - Google is a outstanding source of information and of access to data
  - Google allows to easily download the retrieved information

Google

… and starting sketching a map of communities and stakeholders

---

GL16 Conference, December 8-9, 2014 Library of Congress, Washington D.C., USA

Conclusions & Remarks

The processing has been time-consuming and little fulfilling:
- there is a big amount of disorganised and badly-structured material
- there are many web pages with access denied
- the term "letteratura grigia" often appears simply as a <u>description</u> of the material but the <u>content</u> is not available (e.g. catalogues, lists of references, glossaries, encyclopaedia entries …)

There is a great variety of non-academic communities, each of them having a different approach to GL material:
- grey documentation does not have any frontier and belongs to the most diverse fields
- grey literature is more miscellaneous and more variously represented

The Web helps in disclosing what is available as well as in pointing out the difficulties which the communities still face in organising the grey material

26

Shades of Grey:
Integrating Metadata for Discovery in a
Mixed-Content Single-Subject Library

Genevieve Podleski
Federal Reserve Bank of St. Louis, United States

In 2004, the Federal Reserve Bank of St. Louis Research Library saw an opportunity to build a digital library focused on economic and banking history and created the Federal Reserve Archival System for Economic Research (FRASER). Beginning with historic economic statistical publications produced by both the federal government and the Federal Reserve Banks, FRASER's digitization projects focused on granting researchers access to historic economic conditions in the United States. To provide further context, we added monetary policy documents, including congressional hearings, speeches and testimonies from Federal Reserve Chairman and Board members and developed deep collections on monetary policy history and early central banking in the United States. Through partnerships with other Federal Reserve Banks, we have digitized institutional publications not widely available to the public. In recent years, we have expanded beyond institutional documents and added significant archival collections relating to economic and banking history.

This mix of grey lit (government documents and the quasi-governmental output of the Federal Reserve), white lit (published books and periodicals), and archival material posed significant challenges to organization and discovery. Because each type of material had its own cataloging and organizational conventions, the back-end organization of the FRASER website had been set up with discrete and sometimes conflicting metadata schemes. Beginning in 2013, FRASER librarians began work on a coherent metadata plan that would reorganize the digital library under a single, comprehensive metadata and description scheme in order to maximize usability for internal and external researchers and allow for future development of the collection.

Following a review of literature, we looked to digital libraries both in the U.S. and abroad for examples and case studies. Based on this research, we developed a tentative plan using modified Dublin Core. We then created a crosswalk of existing FRASER metadata to the Dublin Core model and selected as test cases. After multiple rounds of testing and evaluation, we discovered that the complexity of our documents required a more flexible approach than was provided by Dublin Core. MODS, based on library cataloging standards, offered a more robust and flexible method of representing and describing our metadata.

The new MODS-based metadata scheme and a controlled vocabulary, which include documentation accessible to both FRASER librarians and clerks who do the majority of the day-to-day work on FRASER, are scheduled to be implemented in mid-2014.

Bionote

Genevieve Podleski holds a master's degree in library and information science from Pratt Institute in New York, NY and is a master's candidate in history at the University of Missouri-St. Louis. She is currently a digital projects librarian with the Federal Reserve Bank of St. Louis, where she works with FRASER (http://fraser.stlouisfed.org), an online library of banking and economic history that is freely available to the public. Her research interests include metadata, cultural informatics, and open scholarship. Email: genevieve.m.podleski@stls.frb.org

SHADES OF GREY: INTEGRATING METADATA FOR DISCOVERY IN A MIXED-CONTENT SINGLE-SUBJECT LIBRARY

GENEVIEVE PODLESKI
FEDERAL RESERVE BANK OF ST. LOUIS

FRASER (fraser.stlouisfed.org) is a mixed subject and institutional repository that preserves and provides access to economic and banking data and policy documents.
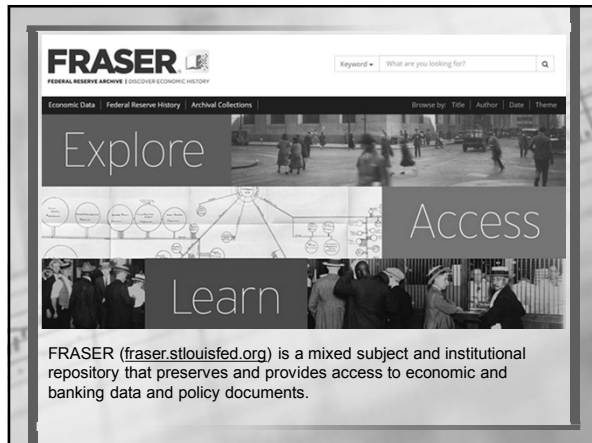
FRASER IS GREY

- **Federal Reserve publications make up approximately 30% of FRASER's nearly 1200 titles, including economic datasets**

- **United States government publications make up more than 50% of FRASER titles**
  - **Legislation**
  - **Hearings**
  - **Reports**
  - **Data**

- **FRASER also holds more than two dozen digitized archival collections**

WHERE DID WE START?

- FRASER began as in 2004 as a data preservation and accessibility project – not a "digital library"
- Metadata was developed ad-hoc
- Each section of the public website had its own metadata set in the database

Publications

+

Archives ➤ **FRASER**

+

Statistical Data
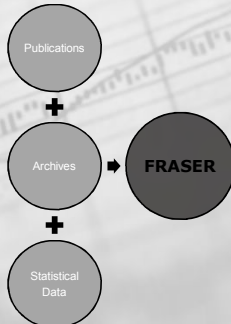
GOALS

- More attractive and browsable public site launched in conjunction with the Federal Reserve centennial
- Better interoperability with other collections and with DPLA
- Improved ability to share content, including dynamically-generated research citations

THE PROCESS

DRAFT

REVISE

DISCUSS

IMPLEMENT

EVALUATE
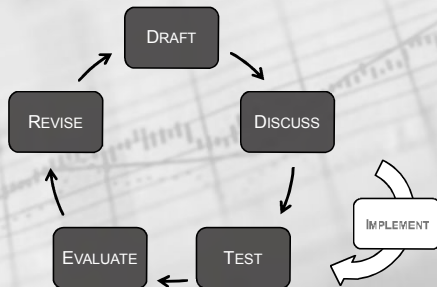
TEST

DUBLIN CORE

- Pros
  - commonly used
  - easy crosswalk
  - DPLA compliant
  - simple and straightforward

- Cons
  - broad instead of precise
  - many element usage conventions didn't fit FRASER collection
  - poor representation of related items and series

TESTING EXPOSES PITFALLS

- How much work will this take to implement?
  - Can data entry be automated?
  - Will you have to create vocabularies and best practices from scratch?
  - How much of the collection will an 'out of the box' implementation describe accurately?

- What's out of our control?
  - Technological limitations
  - Organization of special collections or collections from contributing institutions

MODS

- More 'library-like' – AACR2 and RDA compliant
- Avoids Dublin Core implementation issues
- Allows import/conversion of MARC records
- DPLA compliant

But…

   Still isn't a perfect fit for archival collections, some titles, multiple authors, or non-print items

31

IMPLEMENTATION

- Once a standard is selected, you still have to implement it.
  - Carefully document choices you make when implementing
  - Make adjustments to the standard if necessary
  - Automate what you can
  - Resign yourself to doing the grunt work
  - May need to change materials or practices and not the standard

The perfect is the enemy of the good.
(Le mieux est l'ennemi du bien.)
— **Voltaire**

LESSONS LEARNED

- Begin at the beginning, if possible
- Test, test, test, and then test again
- Even the best fit isn't a perfect fit
- Sometimes you have to change your practice
- One step at a time…
- …even if you step backwards
- Document *everything*
- Just Do It.

QUESTIONS?

Genevieve Podleski
genevieve.m.podleski@stls.frb.org

How Grey Literature Informs Policy and Decision Making: The Necessity to Understand the Processes

Bertrum H. MacDonald, James D. Ross, Suzuette S. Soomai, and Peter G. Wells
Dalhousie University, Canada

Effective advocacy for grey literature must be based on understanding the environments in which it is used. As advances in communications technologies continue to occur at seeming breath-taking pace, all forms of information are being affected. Evolving publication practices are presenting new communication opportunities, in addition to disruptions of established patterns, as long-standing genres are being reshaped by powerful technological and societal changes. Disruptions can cause discomfort and anxiety, but opportunities to promote the value of particular information genres also arise. Grey literature, for example continues to be produced in large quantities, which suggests that its importance in communication may be increasing rather than diminishing. Advocates of grey literature may believe this genre is undervalued or misunderstood, but lobbying for grey literature in the absence of understanding the contexts in which it is or can be used will likely fail unless information activity in those settings is understood. One prominent context encompasses public policy and decision making where grey literature is often present but typically not noticed. Policy and decision-making are notably complex processes and increasing attention is being placed on developing an understanding of the research-policy interface and evidence-based policy making in particular. Conferences (e.g., Science Advice to Governments, Auckland, New Zealand, August 2014), evidence information services (e.g., one launched in the United Kingdom in 2014), research programs and institutes (e.g., Environmental Information: Use and Influence, Dalhousie University), and other initiatives emphasize the importance of understanding the relationship between research and policy, a sometimes contentious and even dysfunctional activity. Drawing on findings from research conducted within the Environmental Information: Use and Influence research program, which involves governmental, intergovernmental, and non-governmental organizations, we will outline roles for grey literature in policy and decision-making contexts. We will note, for example, types of grey literature used in these contexts, we will identify preferences for specific features of useable information by managers and policy makers, and we will outline pathways of research evidence, some of which is produced as grey literature. Information use is a non-trivial phenomenon that must be understood in advance of advocating the value of grey literature.

Bionotes

Bertrum H. MacDonald is a Professor of Information Management in the School of Information Management at Dalhousie University, Halifax, Canada. With a background in science (BSc, Biology), history of science (MA), and information science (MLS, PhD), he pursues research that investigates the dissemination and use of scientific information in historical and contemporary contexts. He pursues interdisciplinary research, particularly within the Environmental Information: Use and Influence initiative (www.eiui.ca), since this work tackles large questions from the point of view of several relevant disciplines. He has been Director of the School of Information Management and Associate Dean (Research) in the Faculty of Management at Dalhousie University. He can be seen speaking about research projects at local, national, and international levels, and he holds executive positions with national and international associations. In 2004, he won the International GreyNet Award with his research colleagues, Ruth Cordes and Peter Wells. He is the recipient of the Marie Tremaine Medal, the highest award of the Bibliographical Society of Canada, and he was awarded a Dibner Research Fellowship at the Smithsonian Institution in 2001. Email: bertrum.macdonald@dal.ca

Bionotes *CONTINUED*

Suzuette S. Soomai is an interdisciplinary doctoral student in the Faculty of Graduate Studies. Her research focuses on the role of fisheries information, published extensively by national and international governmental organisations, in policy making for fisheries management. She holds a Master of Marine Management from Dalhousie University and an MPhil (Zoology/Aquatic Ecology) and BSc from the University of the West Indies, Trinidad and Tobago. Prior to her studies at Dalhousie University, Suzuette was a Fisheries Officer with the Ministry of Agriculture, Land and Marine Resources in Trinidad and Tobago and was a member of fisheries scientific working groups led by the United Nations Food and Agriculture Organisation and the Caribbean Regional Fisheries Mechanism. She has special expertise in fisheries resource and coastal zone management and has completed national and regional fish stock assessments while interacting with a diverse range of stakeholders in the Caribbean. In 2012, she was awarded a major doctoral fellowship from the Social Sciences and Humanities Research Council of Canada. Email: suzuette.soomai@dal.ca

Peter G. Wells is an Adjunct Professor in the School for Resource and Environmental Studies and the Marine Affairs Program, Faculty of Management, and a Senior Research Fellow, International Ocean Institute, at Dalhousie University, Halifax, Canada. He holds a BSc (Biology) from McGill University, a MSc (Zoology) from the University of Toronto, and a PhD (Zoology/Aquatic Toxicology) from the University of Guelph. After over 30 years of public service, he took early retirement from Environment Canada, in 2006, to focus on his research. He concluded his work with Environment Canada as Head, Coastal and Water Science, and Senior Research Scientist, Coastal Ecosystems. He served the United Nations Joint Group of Scientific Experts on Marine Environmental Protection in various capacities for 14 years, and taught an international marine pollution course in Bermuda for 16 years. His current research includes choosing indicators for coastal ecosystem health, utilizing blue mussels for monitoring chemical contaminants, and evaluating the use and influence of marine environmental information in environmental policies and decision making. He was elected Fellow of the American Association for the Advancement of Science, and is a recipient of Dalhousie's highest award for teaching excellence by part-time faculty. Email: oceans2@ns.sympatico.ca

Is the Licensing of Grey Literature Using the Full Palette of "Contractual" Colors?

Tomas A. Lipinski,
School of Information Studies, University of Wisconsin—Milwaukee, United States

Andrea J. Copeland,
School of Informatics and Computing Indiana University, United States

This paper builds upon the analysis used in Tomas A. Lipinski and Andrea Copeland, Look before you License: The Use of Public Sharing Websites in building Patron Initiated Public Library Repositories, 42(4) PRESERVATION, DIGITAL TECHNOLOGY & CULTURE, at 174, November 2013 (pp. 174–198) and to a lesser extent on Tomas A. Lipinski (forthcoming), Click Here to Cloud: End User Issues in Cloud Computing Terms of Service Agreements, in IMCW2013 Proceedings (Yaşar Tonta editor, 2014) (Springer: Communications in Computer and Information Science Series).

Problem/Goal: What terms and conditions are found in licenses or TOS (Terms of Service) governing the use of Grey Literature and are these terms and conditions adequate to preserve, make accessible and use Grey Literature?

Research Method/Procedures: Several licenses or TOS governing Grey Literature collections are reviewed. The licenses identified for consideration were sourced from the affiliations of GL Conference presenters (not including poster presentations) since 2009 (2010-2014 or GL12-GL15) or from the grey literature collection that was a focus of a presentation. Not all presenters were affiliated with collections of grey literature and not all TOS were available through a web search or from the affiliation website of the presenter. This initial review resulted in 10 TOS (OpenGrey (a Creative Commons based license), CNRS the National Center for Scientific Research, CERN the European Organization for Nuclear Research, Alberta Health Services, Research Gate, KISTI the Korea Institute of Science and Technology Information, British Library, Italian Institute of Computational Linguistics, Library of Zeeland (Netherlands), and the National Documentation Center of the National Hellenic Research Foundation). Where a country offered multiple grey literature collections, or presenters only one was selected for inclusion in the analysis. In addition, two commercial suppliers (EBesco Green File and Wolters Kluwer OVID) were included for comparison. The licenses fell into three general categories: those collections governed by a Creative Commons license, those from a governmental agency employing TOS or license and those supplied by a commercial vendor and subject to license and contract. The TOS from government entities varied greatly. Representatives of each TOS type are reviewed in detail and deconstructed in terms of legal essence or effect in order to present a "landscape" or descriptive overview of the legal aspects of Grey Literature licenses and TOS in current use. While not all licenses and TOS reviewed fell within U.S. jurisdiction the contract and other legal principles applied are common to all countries. Second, specific terms and conditions were compared in order to identify problematic provisions in light of one possible measure of efficacy; proposing that the following minimum characteristics should be found or addressed in the required Terms of Service (TOS) in public repositories: Functionality, Integrity, Provenance and Permanence (FIPP).

Results: The results will allow creators as well as users of Grey Literature to better understand both the limitations and possibilities of maintenance, access and use of Grey Literature collections. A further application of the FIPP model or framework for TOS assessment is demonstrated. The FIPP model is further modified to provider further usefulness.

Bionotes

Tomas A. Lipinski obtained his J.D. from Marquette University Law School, LL.M. from The John Marshall Law School, and Ph.D. from the University of Illinois at Urbana-Champaign. Professor Lipinski has worked in a variety of library and legal settings including the private, public and non-profit sectors. Professor Lipinski teaches researches and speaks frequently on various topics within the areas of information law and policy, especially copyright, free speech and privacy issues in schools and libraries. In fall of 2005, Professor Lipinski was placed on the Fulbright Senior Specialist Roster and was named a member of the Global Law Faculty, University of Leuven in Fall of 2006. Email: lipinski@sois.uwm.edu

Andrea J. Copeland is an assistant professor in the School of Library and Information Science at Indiana University at Indianapolis. She received a PhD from Drexel University in information studies in September 2009. The digital preservation practice of public library users is the subject of her dissertation. Her studies at Drexel were supported by the Laura Bush 21st Century Librarian Fellowship from the Institute of Museum and Library Services. She was also the recipient of the Society of Woman Geographers Fellowship in 2002 for her research on the variation of public library use in New York City neighborhoods. Email: ajapzon@iupui.edu

A Comparative Analysis of Grey Literature Terms of Use

presented for

GL-16

on December 8, 2014

© Tomas A. Lipinski and Andrea Copeland, 2014
tlipinsk@uwm.edu,

1

Provider: Warranty Disclaimers

- Example: "**As is**…makes **no representations or warranties** with respect to this Website or its content, all warranties relating to this Website and/or its content and/or any website to which it is linked are here to the fullest extent by law excluded…No representations or warranties are given as the **accuracy or completeness** of the information provided…"
  - National Documentation Centre, National Hellenic Research Foundation.

- Purpose: Alerts user that there are no legal guarantees regarding the functioning of the website or its content.

- Advantages: Limits liability of the provider for errors and omissions relating to the content provided and functioning of its website.

2

Provider: Change in Content

- Example: "Changing circumstances may cause STFC to have to **change** the information and **contents** of its pages at any time."
  - STFC (Science and Technology Facilities Council, United Kingdom).

- Purpose: Allows the provider to add, edit or remove content in response to a take-down request or other reason.

- Advantages: Obviates the need to make users aware that content has changed. Material changes in content do not constitute breach of contract.

3

Provider: Take-Down Mechanism

- Example: "Pursuant to Title 17, United States Code, **Section 512(c)(2)**, notification of claimed copyright infringement under United States copyright law should be sent to Service Provider's Designated Agent."
  – Library of Zeeland, Netherlands.

- Purpose: Allows copyright holders a mechanism through which works claimed to be infringing can be removed.

- Advantages: A provider can limit its monetary liability for secondary copyright infringement if it expeditiously removes or disable access to content claimed to be infringing.

4

Provider: Use Equals Assent

- Example: "By **using** the Site, you, the user agree to the following terms and conditions…" "**Please review** the Terms and Conditions **regularly** as your **continued use** of the site will constitute your agreement to any changes...date of the last revision…at top of page."
  – Alberta Health Service.

- Purpose: Access of the website content can bind the user if the terms are made available and it is clear to the user that by proceeding further the user will be bound by the terms and conditions.

- Advantages: All use of the provider's website is made subject to the terms and conditions. Changes in the terms and conditions can also be incorporated into the agreement when a user next uses the website.

5

User: Notice of Term/Content Changes

- Example: Absent! Some vendors do provide notice:
  – ARTstor license "may be amended from time to time." Changes that "*materially conflict*" with existing terms: *30 days notice via electronic form*. If the licensee objects, the licensor will "use all *reasonable efforts to agree* upon *mutually acceptable language*."

- Purpose: To inform the user of changes in the contractual agreement with content providers in order that may select to accept or reject the new terms.

- Advantages: For changes to terms or content to be legally enforceable, users must have the opportunity to assent to those changes.

6

User: Warranty of Non-infringement

- Example: Absent! Some vendors do provide a warranty of non-infringement and indemnification:
  - "Licensor [Greenwood Press] represents and *warrants* that it has the ***right and authority to make Licensed Materials available*** pursuant to these terms and conditions and that providing the Licensed Materials to Licensee **does not infringe** upon any copyrights, patent, trade secret, or other proprietary right of any third person."
- Purpose: A disclaimer of non-infringement means that the content supplied by the licensor does not infringe another's copyright. Under the laws of copyright liability using (printing, downloading, reposting, distributing, etc.) a work that is infringing is an act of infringement, making a user that prints, downloads, reposts, etc. liable as well.
- Advantages: Users can use the content made available without fear of copyright infringement as long as their use is not itself infringing.

7

User: Restoration after Take-down

- Example: "If you believe your own copyrighted material has been removed from the Site in error, you may submit a written Counter Notice to our Designated Agent (as identified above)…"
  - ResearchGate

- Purpose: Allows copyright holders a mechanism through which works removed based on copyright infringement to be restored.

- Advantages: Communication of restoration rights and processes to the contributors, whose content is being removed or disabled, increases the chance of permanent access to items within a provider's collections.

8

User: Rights

- Example: "perform work, or transmit or store data consistent with the stated goals, policies, and conditions of use as defined by the bodies or bodies granting you access" -CERN (European Organization for Nuclear Research)

- Purpose: For users to understand how they may use the content and services of the provider in relation to copyright law, fair use, and any enforceable contract they've agreed to.

- Advantages: Couching terms of use in the form of legal agreements (contractual license) allows for the clarification of unresolved or contentious areas of the copyright law and thereby increasing the scope of use rights.

9

NTK

50°6'14.083"N, 14°23'26.365"E

Národní technická knihovna
National Technical Library

NU¶ŠL

National Repository in the Czech Republic of Grey Literature
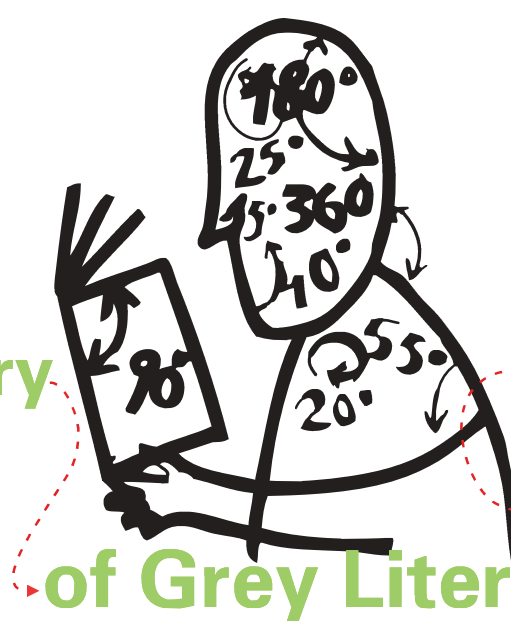
Features

Website: www.nusl.cz
Provider: National Technical Library
Records: over 200 000 records
Partners: over 90 organizations
Source area: Academy of Science, Public Research Institutions, Universities, Libraries etc.
International Cooperation: OpenGrey, DRIVER, ROAR, OpenDOAR
Collection provenance: Czech Republic

Based on

Project: The Digital Library for Grey Literature
– Functional Model and Pilot Implementation (2008 – 2011)
Participants: the National Technical Library, the University of Economics Prague
Financial support: by the Ministry of Culture of the Czech Republic acknowledged

Goals

- Central access to grey literature and the results of research and development in the CR
- Support of science, research and education
- Systematic collection of metadata and digital documents
- Long-term archiving and preservation
- Cooperation with foreign repository

Support of expert discussion about Grey Literature

Annual Workshops:
http://nrgl.techlib.cz/index.php/Workshop
Informative Web pages: **http://nrgl.techlib.cz**
Publication: Grey Literature Repositories
http://nrgl.techlib.cz/index.php/Book

Number of records in NUSL

150000
120000
90000
60000
30000
0

2010   2011   03/2012
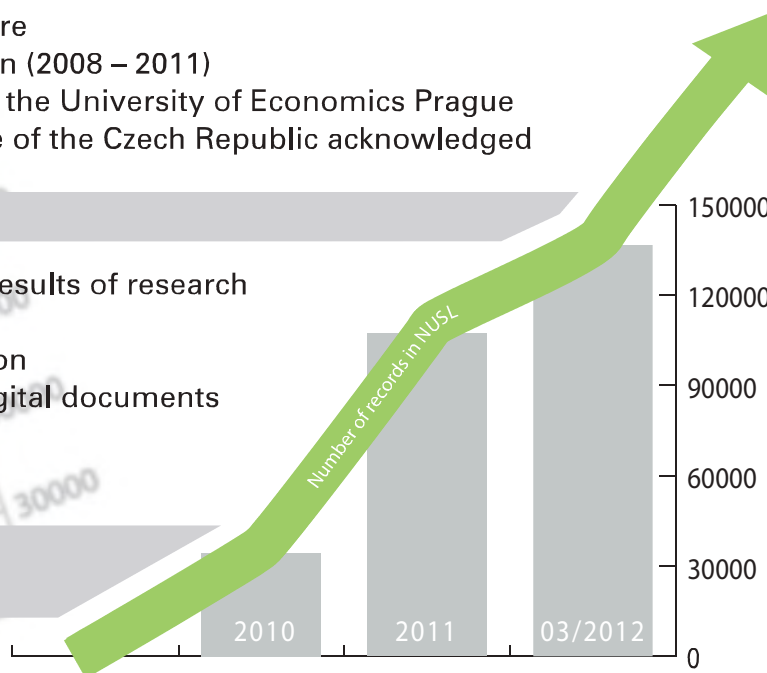
www.nusl.cz

Publishing Geodesy, Topography and Cartography Research via Invenio

Jiří Drozda and **Veronika Synková,**
Research Institute of Geodesy, Topography and Cartography;
Petra Pejšová, National Library of Technology, Czech Republic

We would like to present open source system Invenio that helps to really small specialized library to provide high standard services and grey literature open the access for limited budget and minimal staffing. It will describe access to research outputs in branch of geodesy, topography and cartography. The Surveying Library is unique library focused on geographic branch in the Czech Republic and surrounding countries. The Surveying Library keeps more than 44 thousands units/books, some of them more than 400 years old. Although the Czech Republic is a small country the community of geodesy, topography and cartography is not concentrated in one place. There are four research, ten administrative workplaces and seven universities with geographic or geodetic department. In addition the Surveying Library is located in small city Zdiby outside of Prague and can lure more travelers than students or researchers who should be main users of library. The Surveying Library found the way how to take advantage of the time when readers do not use their legs but mainly their fingers on the keyboard to reach documents they are looking for.

Information technology, open source software and the common availability of the internet allowed that the library can provide their services remotely and at higher quality standards than before. In the past the information sources were fragmented. Users had to know how to reach library catalog, the digital library, electronic resources, archives; many documents were hidden only on the webpages especially grey literature (conference proceedings, research reports, etc.). Today all information sources are provided from one common location and are part of one system in institutional digital repository. It is managed through the implementation of open source software Invenio, whose functions will be presented on the example of implementation in our Surveying Library.

Bionotes

Jiří Drozda is currently chief of Information Department and Surveyors Library of Czech Research Institute of Geodesy, Topography and Cartography (VUGTK). He graduated in Geodesy and Cartography in 1984 at Military University in Brno (CR), special courses in 1997 at Defense Mapping School in Fort Belvoir (USA) and in 2008 at NATO Communication and Information Systems School in Latina (IT). Before coming to VUGTK, he worked for Military Geographic Institute as deputy director (1996-2003) and for NATO as International Civilian Consultant for geographic information support in Balkan territory. He is involved in various national and EU projects focused to education and dissemination of geographic information. Email: jiri.drozda@vugrk.cz

Petra Pejšová graduated in Information science and librarianship at the Charles University in Prague. Currently she is a manager of the National repository of Grey Literature in the Czech Republic. Its goal is a systematic collection, long-term archival and provision of access to specialized grey literature, especially to do with research and development, civil service and education, as well as from the business sphere and "open access" at the national level. She works in the National Library of Technology as a head of the Digital National Technical Library. Through lectures and publications she provides information on grey literature in the Czech Republic. She is a member of Czech affiliate team for dissemination free licenses Creative Commons. Email: petra.pejsova@techlib.cz

Veronika Synková studied information science and librarianship at the Institute of Information Studies and Librarianship at the Faculty of Arts of Charles University in Prague, Czech Republic. She works as a system librarian in the Library of Antonin Svehla of the Institute of Agricultural Economics and Information. She is a member of the team responsible for the library system Aleph and the discovery system Primo. She cooperates on implementation of the system Invenio with the Surveying Library of the Research Institute of Geodesy, Topography and Cartography.
Email: veronika.synkova@gmail.com

GL 16, Washington D.C., December 8, 2014 8, 2014

Publishing Geodesy, Topography and Cartography
Research via Invenio

Jiří Drozda, Veronika Synková

**Research institute of Geodesy, Topography
and Cartography
Zdiby, Czech Republic**

Petra Pejšová

**National Library of Technology
Prague, Czech Republic**

NTK
Národní technická knihovna
National Library of Technology

Research Institutions in the Czech Republic

The Academy of Sciences of the Czech Republic

61 research institutes

Academy of Sciences Library - Central Dig. Rep.

Public Research Institutions

Lack of technology and methodological support

Comercial Research Institutions

Suficient financial funds

Library systems

- Comercial - Aleph, KPWin, Clavius…
- Non-comercial - Evergreen, Invenio

Digital Repositories

- Comercial -Theses, DigiTool, STAG, ARL …
- Non-comercial - DSpace, Invenio

Surveying Library of VÚGTK

- **branch specialized library with an unique library fund focused on all geo-informatics not only in the Czech Republic**

BUT

- **the library, which a few years ago was threatened to extinction** – outdated library system, fragmented information on the web, library services requires a personal visit, located in small town = low visit frequency

THAN

- **the library starts to transform into a modern information centre, focused mainly on on-line services**

Fragmented Resources
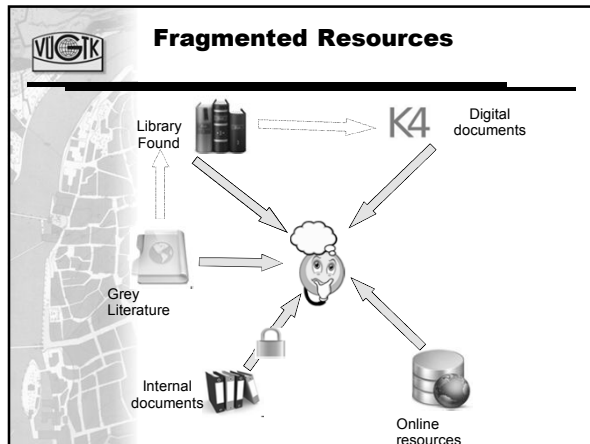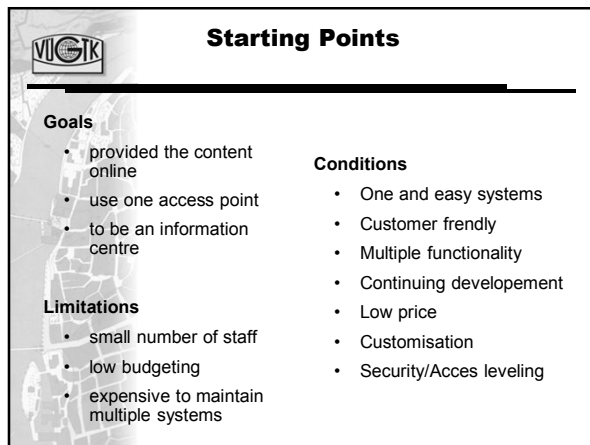
Starting Points

Goals

- provided the content online
- use one access point
- to be an information centre

Limitations

- small number of staff
- low budgeting
- expensive to maintain multiple systems

Conditions

- One and easy systems
- Customer frendly
- Multiple functionality
- Continuing developement
- Low price
- Customisation
- Security/Acces leveling
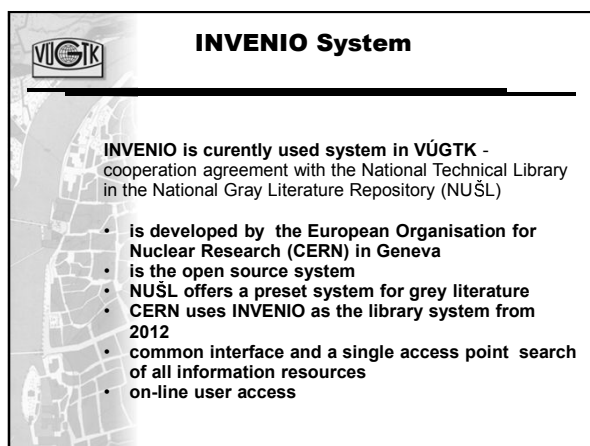
INVENIO System

INVENIO is curently used system in VÚGTK - cooperation agreement with the National Technical Library in the National Gray Literature Repository (NUŠL)

- **is developed by the European Organisation for Nuclear Research (CERN) in Geneva**
- **is the open source system**
- **NUŠL offers a preset system for grey literature**
- **CERN uses INVENIO as the library system from 2012**
- **common interface and a single access point search of all information resources**
- **on-line user access**

The resources of the library in the INVENIO system
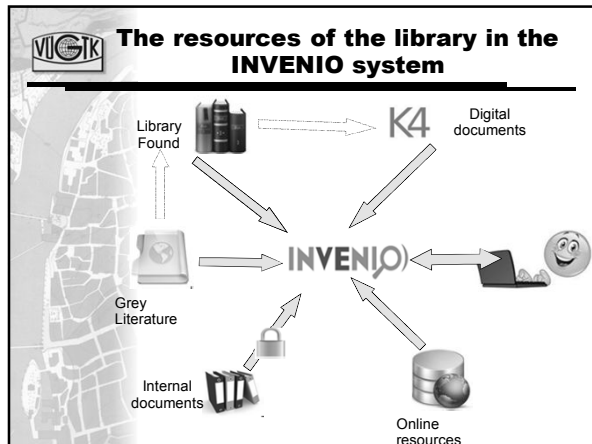
INVENIO Implementation

Integrated Funds
- Library fund
- Digital library fund
- On-line access to scientific journals
- Grey literature
- Internal documents

Library system functions
- Circulation services
- On-line user accounts
- History search
- Alert services
- User thematic baskets

Custom collections/folders in INVENIO

knihovna.vugtk.cz

Record in INVENIO with link in to digital library Kramerius

User account

Plans for Future

- Beta debugging 1Q/2015
- Storage of scientific data
- Automatic harvesting by NUŠL and other systems
- Support and training of scientists to use INVENIO system
- Creating of new geo-thematic collections

Free Licences and Creative Commons: A Powerful Tool for Open Access Publishing in Grey Literature

Petra Pejšová, National Library of Technology, Czech Republic
Marcus Vaska, Knowledge Resource Service, University of Calgary, Canada

In today's increasingly technologically savvy information society, "using remote access and free content to open doors for science students", a statement made by NANSLO Lab director Daniel Branan (www.scoop.it/t/ava-openeducation), is yet another example of ongoing efforts to make information more openly and freely available and accessible. Although Branan focused his remarks on the scientific community, this applies to more than one specific subject field. Rather, scientists, teachers, artists, sociologists, programmers, as well as professionals from the arts industry and economics are increasingly becoming involved in sharing and reusing their work. Open content provides an opportunity to shorten the time for research to become available, not repeat research already conducted, have data to compare, collect background information for a project, and numerous other possibilities.

Despite the well-intentioned mandate of a Creative Commons license, the free distribution of an author's work is still "governed by applicable copyright law." (Wikipedia, n.d.) Jack Andraka, an advocate for the Open Access Movement, laments the disappointment that can occur due to publication and distribution restrictions: "I've seen so many great ideas get killed in the lab when my peers are stopped by closed access [to research articles]" (http://teamopen.cc/jack/). Open licencing is a strong instrument ensuring open access to research data.

Research Method/Procedure:

This project will uncover open licenses and describe how they are used, focusing on Creative Commons free licences, the most widely known worldwide. The Open Access movement has begun gaining greater acceptance, with numerous institutions either strongly encouraging and/or requiring their faculty, students, and staff to deposit their scholarly work in the institutional repository. As a case in point, the University of Liege in Belgium established a mandate in 2008 whereby all publications must be deposited, including the full text of articles „as soon as the article is accepted by the editor" (http://www.openaccessmap.org/list/) Despite the well-intentioned means of encouraging authors to deposit their works in the public domain via open-content licenses, controversy still remains that this act can alter the original author's ownership, particularly since "all transfers or licenses of copyright interests by a work's author are revocable." (Armstrong, 2010, p. 360). The University of Liege has countered this argument with their ORBi (Open Repository and Bibliography) open access repository; a clause has been added stating that access to an author's full text articles „will only be granted with the author's consent and according to the rules applicable to author's rights and copyrights" (http://www.openaccessmap.org/list). This increased visiblity in publications and access to research has resulted in ORBi currently holding a ranking of 34 out of 1746 repositories worldwide, recording more than 2 million downloads since its inception (http://orbi.ulg.ac)

Via a survey, international, national, subject, and institutional repositories will be selected, in order to determine if Creative Commons licences are being used at these facilities and if so, how and in what way (i.e. which type of documents are being deposited?, what is the degree of usage? etc.). The survey will focus on the different Creative Commons licenses available, and how this affects open access and copyright restrictions.

Results:

We believe that results gleamed from the survey will not only provide us with a comparative environmental scan of the existence of Creative Commons licenses at various institutions, but will also reveal insufficiencies and recommend approaches on how to increase the use of these licences in grey literature repositories. It is anticipated that this venture will generate renewed interest and awareness in creating a more seamless link between open access publishing and grey literature. It is in this research context that the technology and innovation triangles combine, "extending the scope beyond R&D [research and development]" (Pant and Hambly-Odame, 2010), to the grey literature community as a whole. While certain document types may never be deposited into an institutional repository, and some authors may voice concerns about feeling obligated to adhere to such a mandate, the benefits clearly outweigh any potential harms. Open Access publishing in the grey literature domain via the use of free Creative Commons licenses creates the multiplier effect, "permitting the creation of new works which may never have come into existence" (Armstrong, 2010, p. 368).

References

Armstrong, T.K. (2010). Shrinking the Commons: Termination of Copyright Licenses and Tranfers for the Benefit of the Public. *Harvard Journal on Legislation, 47*(2), 359-423.

Charting the Growth of and Development of Open Access Globally (n.d.). Retrieved April 3, 2014 from http://www.openaccessmap.org/list

Open Repository and Bibliography (ORBi). (2014). Retrieved April 3, 2014 from http://orbi.ulg.ac.be/
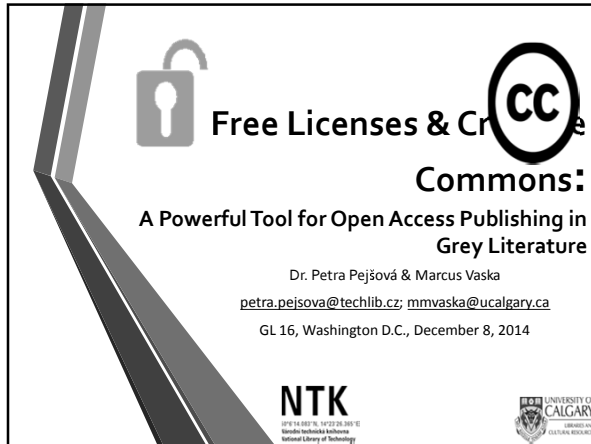
Pant, L.P., and Hambly-Odame, H. (2010). Creative Commons: Non-Proprietary Innovation Triangle in International Agricultural and Rural Development Partnerships. *The Innovation Journal: The Public Sector Innovation Journal, 15*(2). Retrieved March 5, 2014 from http://www.innovation.cc/scholarly-style/pant_odame_creative_commons4final2rev.pdf

Wikipedia (n.d.) *Creative Commons License.* Retrieved March 5, 2014 from http://en.wikipedia.org/wiki/Creative_Commons_license

Bionotes

Petra Pejsova graduated in Information science and librarianship at the Charles University in Prague. Currently she is a manager of the National repository of Grey Literature in the Czech Republic. Its goal is a systematic collection, long-term archival and provision of access to specialized grey literature, especially to do with research and development, civil service and education, as well as from the business sphere and "open access" at the national level. She works in the National Library of Technology as a head of the Digital National Technical Library. Through lectures and publications she provides information on grey literature in the Czech Republic. She is a member of Czech affiliate team for dissemination free licenses Creative Commons. Email: petra.pejsova@techlib.cz

Marcus Vaska is a librarian with the Knowledge Resource Service (KRS), University of Calgary, responsible for providing research and information support to staff affiliated with an Alberta Cancer Care research facility. A firm supporter of embedded librarianship, Marcus engages himself in numerous activities, including instruction and research consultation, with numerous research teams. An advocate of the Open Access Movement, Marcus' current interests focus on creating greater awareness of the importance of creative commons licenses in the publishing of grey literature material. Email: mmvaska@ucalgary.ca

Free Licenses & Creative Commons:

A Powerful Tool for Open Access Publishing in Grey Literature

Dr. Petra Pejšová & Marcus Vaska

petra.pejsova@techlib.cz; mmvaska@ucalgary.ca

GL 16, Washington D.C., December 8, 2014

NTK
Národní technická knihovna
National Library of Technology

UNIVERSITY OF CALGARY
LIBRARIES AND CULTURAL RESOURCES
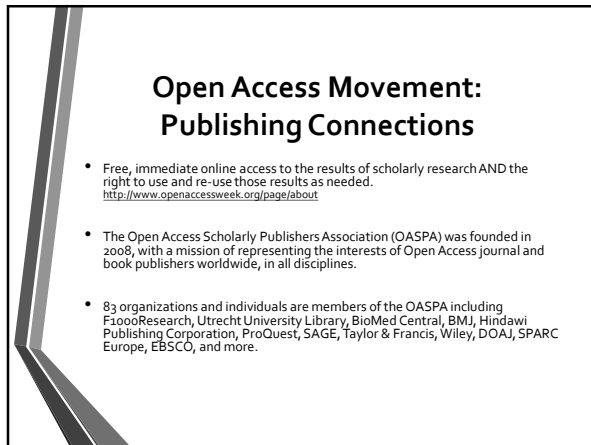
Open Access Movement:
Publishing Connections

- Free, immediate online access to the results of scholarly research AND the right to use and re-use those results as needed.
  http://www.openaccessweek.org/page/about

- The Open Access Scholarly Publishers Association (OASPA) was founded in 2008, with a mission of representing the interests of Open Access journal and book publishers worldwide, in all disciplines.

- 83 organizations and individuals are members of the OASPA including F1000Research, Utrecht University Library, BioMed Central, BMJ, Hindawi Publishing Corporation, ProQuest, SAGE, Taylor & Francis, Wiley, DOAJ, SPARC Europe, EBSCO, and more.

"I've seen so many great ideas get killed in the lab when my peers are stopped by [reading access to the articles]"

http://teamopen.cc/all/

Open Content + Open Licensing = Open Access

TEAM OPEN — Jack Andraka
Fighting cancer with free research.

TEAM OPEN — Sulyo Polyakov

TEAM OPEN — Khalid Albaih

TEAM OPEN — Miyoung Ti

49

Open Access Publishing and Grey Literature

- Grey Literature subject to internal quality assessment (publishing institution's name and reputation at stake).
- No obligation for long-term archiving institution with grey literature.
- Publicly funded grey literature repository (i.e. GreyGuide).
- Copyright exceptions (pertaining to a free use, share-alike license).
- "Grey literature provides an essential complement to peer-reviewed findings." Without open access publishing, locating and retrieving this material would be a daunting task.

Creative Commons Licenses

- One of **7** main licenses permitting free distribution of an otherwise copyrighted work:
  - 1. CC0 No Rights Reserved
  - 2. CC BY Attribution
  - 3. CC BY SA Attribution ShareAlike
  - 4. CC BY ND Attribution – No Derivatives
  - 5. CC BY NC – Non Commercial
  - 6. CC BY NC SA Attribution Non-Commercial ShareAlike
  - 7. CC BY NC ND Attribution Non-Commercial No Derivatives

Creative Commons & Institutional Responses Survey
Where are our Respondents From?

Creative Commons Licenses in Institutional Repositories Worldwide

Are Creative Commons licenses being used at your institutional repository?

84% Yes!

Does your institution have a Creative Commons license policy? If yes, please briefly describe below.

71 % NO

Under Which Terms is a CC License Deed Issued at your Institution?

CC0 No Rights Reserved...	16.22%
CC BY Attribut...	70.27%
CC BY SA Attributi...	56.76%
CC BY ND Attributi...	48.65%
CC BY NC Attributi...	56.76%
CC BY NC SA Attribution...	48.65%
CC BY NC ND Attribution...	59.46%

What Types of Documents are Being Deposited at your Institution via Use of a CC License?

Text	92.11%
Data	36.84%
Video	34.21%
Music/Audio records	21.05%
Pictures	39.47%

51

Besides Creative Commons Licenses, does your Repository Use any Other Freely Available Licenses?

- GNU / GPL / AGPL / Free Documentation License
- ODbL – Open Database License
- Public Domain Mark
- UK Open Government License
- Free ART License
- Non –Exclusive Distribution License
- Metadata Open License

29.55%

70.45%

Creative Commons Licenses Testimonials from our Respondents

- "We don't have much experience and information about 'safe' usage of Creative Commons licenses."
- "We are afraid of not using them properly."
- "Easy to use, general, and validated framework, replies to our need (we do not want others to modify our reports)."
- "Crediting by users is sporadic at best."
- "Researchers are not familiar with licenses."
- "I fully endorse CC licenses and encourage people to use them in my work"

References

- Banks, M. (2004). Connections between open access publishing and access to gray literature. *Journal of the Medical Library Association 92*(2), 164-166.
- Beall, J. (2013). The Open-Access Movement is not really about open access. *TripleC 11*(2), 589-597.
- Myska, M. (2012). A model framework for publishing grey literature in open access. *JIPITEC 4*, 104-115.
- Open Access Scholarly Publishers Association (OASPA). *The International Community of Open Access Publishers.* Retrieved October 8, 2014 from http://oaspa.org
- Suber, P. (2013). *Open Access Overview.* Retrieved October 8, 2014 from http://legacy.earlham.edu/~peters/fos/overview.htm
- Suber, P. (2009). *Timeline of the Open Access Movement.* Retrieved October 8, 2014 from http://legacy.earlham.edu/~peters/fos/timeline.htm
- Surl, L. (n.d.). *Team Open.* Retrieved October 30, 2014 from http://teamopen.cc/all/
- Wikipedia (n.d.) *Creative Commons License.* Retrieved October 8, 2014 from http://wn.Wikipedia.org/wiki/Creative_Commons_license

Degrees of Openness:
Grey Literature in Institutional Repositories

Joachim Schöpfel, Charles de Gaulle University Lille 3
Hélène Prost, CNRS, Associate member GERiiCO, France

The open access principle requires that scientific information be made widely and readily available to society. Defined in 2003 as a "comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community", open access implies that content be openly accessible and this needs the active commitment of each and every individual producer of scientific knowledge.

In spite of the growing success of the open access initiative, a significant part of scientific and technical information remains unavailable on the web or circulates with restrictions. Even in institutional repositories created to disseminate the scientific production of an academic institution, broad and open access to more or less important sectors of the scientific production is restricted. This is because of lack of awareness, embargo, deposit of metadata without full text, confidential content etc.

In order to provide new empirical evidence, 25 large institutional repositories from different continents were selected in the international directory OpenDOAR. For each repository, the access to the full text for different document types was evaluated, and the statistics were analyzed for each site and cumulated.

In the past, we put forward that for grey literature, "open is not enough", i.e., institutional repositories need a set of minimum requirements for grey items such as metadata, selection procedures, quality standards, collection management and clear deposit policy (GL13). We then narrowed our research on electronic theses, recommended five ways how to add value to theses in open archives (GL14) and identified access restrictions to theses in institutional repositories (GL15). This means, while "open is not enough", that all deposited theses in repositories are not open whatever; or more specifically, they are available with different degrees of openness, and some are not available at all.

This year, we return to a larger perspective. Building on our past work and new empirical data from large institutional repositories on different continents, we distinguish between different degrees of openness. Which are the main reasons, which are the stabilizing functions of this situation? The communication tries to provide some elements of understanding, together with good practices and recommendations.

Bionotes

Joachim Schöpfel is Lecturer of Library and Information Sciences at the University of Lille 3 (France), Director of the French Digitization Centre for PhD theses (ANRT) and member of the GERiiCO research laboratory. He was Manager of the INIST (CNRS) scientific library from 1999 to 2008. He teaches Library Marketing, Auditing, Intellectual Property and Information Science. His research interests are scientific information and communication, especially Open Access and Grey Literature. Email: joachim.schopfel@univ-lille3.fr

Hélène Prost is information professional at the Institute of Scientific and Technical Information (CNRS) and associate member of the GERiiCO research laboratory (University of Lille 3). She is interested in empirical library and information sciences and statistical data analysis. She participates in research projects on evaluation of collections, document delivery, usage analysis, grey literature and open access, and she is author of several publications. Email: prost@inist.fr

Degree of openness : Grey literature in institutional repositories

*J
lène Prost*

Université de Lille 3 SCIENCES HUMAINES ET SOCIALES

Gériico

CNrS

16th International Conference on Grey Literature
2014, Dec 8-9 , Washington

% institutional repositories with specific document types

The survey sample

→ Criteria in OpenDoar

Repository type:
Institutional

Content type:
PhD theses and articles (at least)

Size:
10,000+ items (preferred)
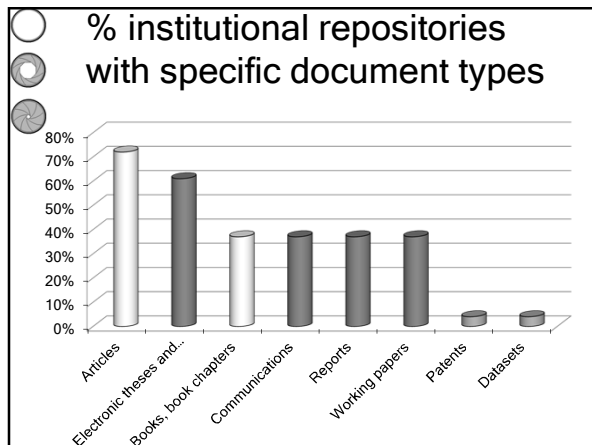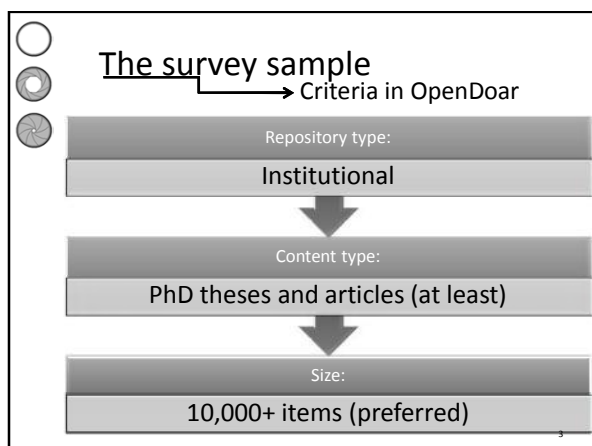
The survey sample

→ 25 selected repositories
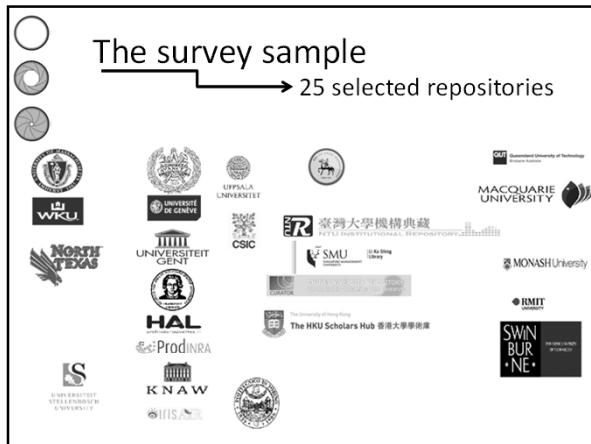
Results

→ Openness per document type

Document type	Number of items	Degree of openness
Communications	490,442	0.21
Theses and dissertations	156,546	0.78
Reports	91,069	0.63
Working papers	3,118	0.96

Results

→ Openness per document type
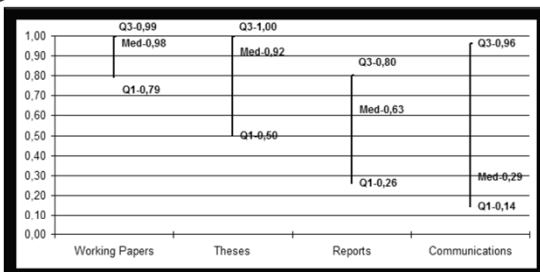
Results

→ Different degrees of openness

- 7 repositories
- more than 80% items with full text in each category

High level

- 6 repositories
- Prefere exhaustiveness to openness

Low level

- 12 repositories
- Theses and working papers more open

In between

☐ Type 2   ■ Type 3   ☐ Type 1

Working Papers   Theses   Reports   Communications
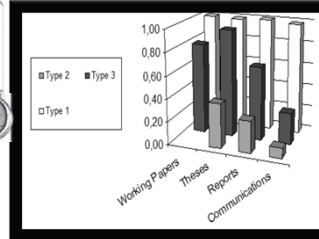
7

Results

→ Grey and white literature

$R^2 = 0,8483$

Grey / White
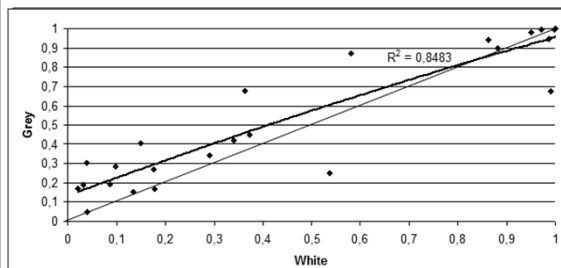
Results

→ Grey and white literature without communications

$R^2 = 0,3553$

Grey (without conferences) / White
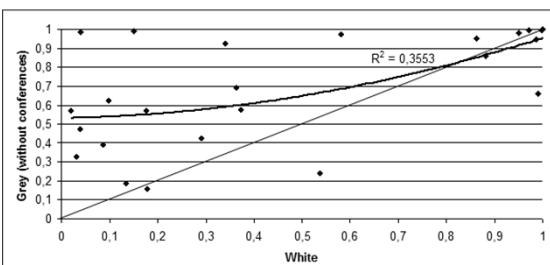
Restricted access

→ Reasons

Communications
- Rights transferred to publishers
- In conference proceedings, special journal issues or monographs
- Significant number under embargo

PhD theses
- Intention to publish as a book

Hosting organizations
- IR is a tool for evaluation
- More metadata than full text

10

Best practices

Commitment to open access principles

Mandatory policies

Selection of deposits

Specific approach for different document types

Institutional workflows

11

Recommandations

→ To improve access to GL

Typology
Metadata should clearly index the document type

Discovery
Allow browsing and searching with document types

Access rights
Indicate the availability or access restrictions for each document

Policy
Foster the deposit of metadata with full text

Differentiation
Distinguish between the different document types

12

Analysis of Collection and management of the Korea National R&D Report

Kiseok Choi, Cheol-Joo Chae, Yong-hee Yae, and Yong Ju Shin,
Korea Institute of Science and Technology Information, KISTI, Korea

The term 'National R&D Program' has been widely used in a Korean society, but it is quite new as a legal term. Recently, its legal concept from the perspective of the unified semantics was defined, and the regulation on its scope has been established. In particular, 'National R&D Program' was officially used as a legal term for the first time in 'Special Act on Innovation in Science and Technology (No. 5340) enacted in April 1997. In the Regulation on the Management of National R&D Programs, 'National R&D Program' is defined in Article 2 as follows: "An R&D program in science and technology, in which a certain R&D project is chosen and partially or wholly funded by the government or public organization in accordance with the related law. It refers to all programs except for the basic programs promoted by government-funded research institutes in accordance with Paragraph 1 of Article 2 of Act on the Establishment, Operation and Fostering of Government-funded Research Institutes". In particular, because the national R&D report, one of intangible research outcomes among national R&D programs, contains important matters on R&D such as R&D method and results, it is the final outcome of the national R&D program regardless of its type. Since it is the evaluation standard of the R&D project and major research outcome as the final outcome which takes place at the end of the R&D program, the systematic management of related information is especially important. In this sense, the establishment of the information distribution system which makes it possible to collect, manage and use national R&D reports has been perceived as one of the most important R&D policies in any country. In fact, each country has systematically managed related duties through a national science & technology information agency. After all, the management of national R&D reports is a critical part in the management of the research outcomes of national R&D programs. This study examines the trend of national R&D report production, progress & assignment of national R&D programs and current non-disclosure reports, investigates the current collection and management of national R&D reports by the KISTI, a professional R&D report management agency in Korea, and proposes a future direction. In addition, the development of an information distribution system and database for national R&D reports and online services are analyzed.
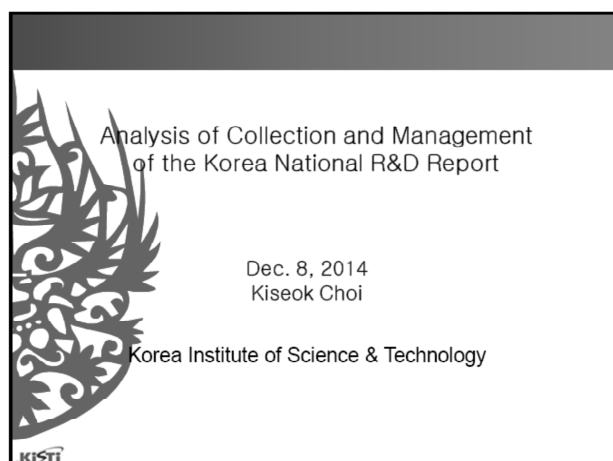
CONTENTS

Introduction

Management Scheme of National R&D Programs

Definition of National R&D Program
- a R&D program in science and technology, in which a certain R&D project is chosen and partially or wholly funded by the government or public organization

Establish" a general distribution system on the related information
- National Science & Technology Information System

Kept and managed by the government
- regulated to be registered on the research outcome management agency(KISTI)

2

The production type and management of national R&D outcomes

Production type of national R&D outcomes

- **The outcome of national R&D programs**
  - all research results which occur after the fulfillment of the research projects regardless of their type and period
  - The outcome types of national R&D programs
    - Tangible/Intangible outcome
  - science and technology research outcome
    - Primary/Secondary outcome

3

The production type and management of national R&D outcomes

Production type of national R&D outcomes

- The outcome types of national R&D programs
  - **Tangible outcome**
    - **various results obtained from the byproducts of R&D, including biological resources, compounds and information data**
  - **Intangible outcome**
    - **knowledge performance acquired through R&D, including knowledge assets (ex: patent, etc.)**
- science and technology research outcome
  - **Primary outcome:**
    - **the result which directly occurs for a short period time**
    - **EX) knowledge creation, technology development, product production and process improvement**
  - **Secondary outcome:**
    - **social and economic effects derived from the primary outcome**
    - **cost reduction, increase in sales and other social and economic contributions**

KISTI   4

The production type and management of national R&D outcomes

Management of the outcome of national R&D programs

- **the management of national R&D reports is a critical part in the management of the research outcomes of national R&D programs**
  - the establishment of the information distribution system for collect, manage and use national R&D reports
  - prevent redundant research or investment and select a new research project through the mutual exchange of R&D information

KISTI   5

Collection and management of national R&D report information

Progress and projects of national R&D programs

- **Current Status of National R&D Programs, their Investments and Specific Projects by Year**
  - **Korea's national R&D programs have continuously risen since the 1990s**
  - **grown by more than 10% annually**

Year Category	2008	2009	2010	2011	2012
Investment (x KRW 100 million)	109,936	124,145	136,827	148,528	159,064
No. of Programs	486	474	483	493	537
No. of Specific Projects	37,449	39,565	39,254	41,619	43,192

KISTI   6

Collection and management of national R&D report information

National R&D report production trend

National R&D Report Production Trend
- Since a strict survey on the number of final report started in 2012, it has declined

Year Category	2010	2011	2012	Total
No. of Report Produced	14,565	15,044	9,646	39,255

7

Collection and management of national R&D report information

Current non-disclosure of national R&D reports by bureau-1

Classification of Non-disclosure Projects of National R&D Programs
- The reports' non-disclosure rate has been as high as 36.9%(last 3 years)

Category	2010	2011	2012	Total (Average)
Total Projects	8,540	8,461	4,805	21,806
Disclosed Projects	4,399	5,635	3,736	13,770
Non-disclosure Projects	4,141	2,826	1,069	8,036
Non-disclosure Ratio (%)	48.5%	33.4%	22.2%	36.9%

8

Collection and management of national R&D report information

Current non-disclosure of national R&D reports by bureau-2

Classification of Non-disclosure Projects of National R&D Programs by Bureau
- Small and Medium Business Administration was the highest with 100%
- Even stay non-disclosure after 3years of patent processing time

	No. of Projects	Disclosed	Undisclosed	Percentage(%)
Ministry of Education, Science and Technology	8,622	8,302	320	3.7
Ministry of Knowledge Economy	5,014	1,838	3,176	63.3
Ministry of Agriculture and Forestry	471	327	144	30.6
Ministry of Health and welfare	780	606	174	22.3
Ministry of Environment	388	371	17	4.4
Ministry of Land, Transport and Maritime Affairs	308	209	99	32.1
Small and Medium Business Administration	3,884	-	3,884	100
Rural Development Administration	1,083	1,073	10	0.9
Korea Food and Drug Administration	861	835	26	3.0
Defense Acquisition Program Administration	4	1	3	75.0
Total	21,415	13,562	7,853	36.7

9

Collection and management of national R&D report information

Collection and management of national R&D reports by the KISTI-1

○ **Development of National R&D Report Database by the KISTI**
- **the ratio of government-led collection and management of national R&D reports is low**
- **the percentage of database to the conventional national R&D reports was about 69%**

Year	2008	2009	2010	2011	2012
No. of Productions	17,233	18,160	14,565	15,044	9,646
No. of Database Development Cases	4,623	7,108	8,540	8,461	4,805

10

Collection and management of national R&D report information

Collection and management of national R&D reports by the KISTI-2

○ **Collection of Non-disclosure National R&D Reports by KISTI**
- **the percentage of non-disclosure national R&D reports is very high**
- **non-disclosure reports are accounting for 48.0% of total national R&D reports**

Category	2008	2009	2010	2011	2012	Total
No. of Total Reports Registered	4,623	7,108	8,540	8,461	4,805	33,537
No. Non-disclosure Reports	3,213	4,836	4,141	2,826	1,069	16,085
Percentage of Non-disclosure reports (%)	69.5%	68.0%	48.5%	33.4%	22.2%	48.0%

11

Conclusion

Conclusion

- The information distribution systems for national R&D reports have been managed quite well
- current information management and distribution of national R&D reports are still very poor
- need to develop a system to disclose non-disclosure national R&D reports and distribute them

Future study

- For the integrated registration and management and efficient distribution of national R&D reports
- use of a standardized R&D report form for the systematic and unified submission and registration
- a copyright issue which could be raised during the online services

12

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

Joanna Hooper, University Hospitals Bristol
NHS Foundation Trust, BRI, United Kingdom

"people conducting systematic reviews...need to ensure they search for trials in the grey, as well as published, literature in order to minimise bias" (Hopewell, 2007)

Grey Literature is now recognised by many information experts as a necessary element in the field of Evidence Based Medicine (EBM). This being said GL still encumbers both the researcher and the LIS professional with many of the same problems that it did in the past. Whilst most peer-reviewed evidence is held within password protected sites, accessible only through costly subscriptions or membership of a professional body there is also excellent information, sometimes freely available on the internet, the quality of which is often hard to define in a quick and proficient manner and its reliability difficult to determine.

So what are the repercussions for the Healthcare professional? And what, more specifically, is the role of the NHS Healthcare Librarian in ensuring that GL is made available for medical staff in order to ensure the best level of patient care? Do LIS Professionals systematically include searching for Grey Literature as part of their daily procedures for capturing and gathering evidence and if so what processes do they have in place? Perhaps more importantly what type of information do LIS professionals consider the term Grey Literature to encompass?

These questions will be addressed with particular focus on the context of GL, assessing its quality, its accessibility, accountability, currency and reliability (both in terms of data and lifespan) and within the framework of Evidence-Based Medicine (EBM).

The proposed methodology is a qualitative, mixed methods semi-structured questionnaire and will be based on both a Likert Scale format to measure attitude and opinions, as well as open ended questions to gather as much rich, narrative data as possible.

Bionote

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

55 Years of the National Union Catalog of Manuscript Collections, NUCMC : A Program of the Library of Congress

Beth Davis-Brown, Peter Goodman, Roberto Sicre, Library of Congress;
Cooperative and Instructional Programs Division, NUCMC Program, United States

The National Union Catalog of Manuscript Collections (NUCMC) program of the Library of Congress provides and promotes bibliographic access to the nation's documentary heritage. This free-of-charge cooperative cataloging program creates online records in OCLC WorldCat on behalf of eligible archival repositories throughout the United States and its territories. During its 55 years of existence, NUCMC has cataloged and made available via OCLC over 130,000 collections for over 1,808 repositories since 1959. Benefits for participants in NUCMC include:

- Researchers worldwide can find your unique, one-of-a-kind collections
- NUCMC records may have useful information about your collections, which can augment their value
- Better descriptive access frequently translates into better security for your collections
- Potential donors may want to give their collections to repositories with complementary collections

Repositories that qualify are unable to contribute to cataloging to OCLC WorldCat and lack an archival cataloger. They must be open on a regular basis to researchers and located in the U.S. or its territories.

An Attempt to Nuance the Understanding of Professional Reports in Archaeology

Lisa Börjesson, Department of Archives, Libraries, Museums;
Uppsala University, Sweden

Professional (i.e. extra-academic) contract archaeology is an internationally widespread practice which contributes significantly to the volume of the archaeology literature corpus. However, professional knowledge production in archaeology, and most notably the professional report genre, is at times described as problematic. The problem descriptions are ambiguous and can be grouped under at least three different topics: concerns for formal quality and technical accessibility, concerns for the comparably low degree of analytical and theoretical synthesizing in reports, and concerns for a lack of recognition of professional archaeology's knowledge production. Technical issues of access are to an increasing extent being solved. Structure standardizations are also developing. Hence the genre becomes more accessible, and the content more readable and informative. However attitudes toward the genre seem to, based on articulations in archaeology text books and journal articles, remain focused on the genre's problems. The aim of my ongoing dissertation research is to nuance the understanding of the professional report genre in archaeology through an analysis of factors shaping the reporting. I argue an improved genre understanding is crucial to diminish cultural issues of access, and also as grounds to manage development of reporting practices.

The aim of this poster is to explicate the mixed methods research design of my ongoing study and to open up for discussions about methodologies for researching contextual factors shaping grey literature. I explore (1) perceptions about grey literature and (2) information policy regulating grey literature using a text centered idea analysis. To contrast the descriptions of grey literature in archaeology literature and the heritage policies' articulations about what the report literature should be, I explore (3) report writers' frames of reference using a quantitative bibliographic method and (4) how reports writers interpret their task and make sense of the report production activity through qualitative focus group interviews. The poster will briefly report on findings from study no. 3, and on preliminary results from study no 2.

Bionote

Lisa Börjesson is a doctoral student at the Department of archival science, library & information science, and museum & heritage studies at Uppsala University, Sweden. She also holds a position as a visiting student researcher at UC Berkeley School of Information 2014-2015. Lisa has a background as a library development consultant working within regional government. She currently works together with archaeologists and information scholars in the research project Archaeological Information in the Digital Society (ARKDIS) funded by the Swedish Research Council, and is a member of the northwestern European Centre for Digital Heritage. Her ongoing dissertation research (2013-2017) explores three contextual factors shaping reporting in professional archaeology: information policy, report writers' frames of reference and their interpretations of the report writing as a professional task. Her research complements initiatives to improve technical access to archaeological grey literature by providing a more nuanced understanding for the report genre.
Email: lisa.borjesson@abm.uu.se

65

An Attempt to Nuance

the Understanding of

Professional Reports in Archaeology

– A Dissertation Research Project

Lisa Börjesson
Department of ALM Uppsala University (SWE)
Lisa.borjesson@abm.uu.se

1

Background I: Contract archaeology

- Most archaeological surveys and excavations are undertaken prior to land development, as contract archaeology.

- Contract archaeology reports make up a significant part of the corpus of archaeology literature. (Aitchison 2010)

2

Background II: Contract archaeology and knowledge (in Sweden)

- Contract archaeology surveys shall document findings, collect objects and disseminate the results of surveys.
  [*Heritage Conservation Act (1988:950), revised up to SFS 2014:649*]

- Documentation shall be done in accordance with scientific standards and aim at creating "meaningful knowledge" (my transl.).
  [Riksantikvarieämbetet (The National Heritage Board) 2012]

3

What is grey literature in archaeology?

"Archaeological reports with limited distribution, usually reports prepared by archaeological contractors." (Darvill 2009)

4

Which are the grey literature problems in archaeology?

a) **Reports are technically inaccessible** (as discussed by e.g. Aitchison 2009; Hardman 2009; McGowan 2009; Stock and Sampité 2009)

b) **The quality of report content is not satisfying** (as discussed by e.g. Bahn 2012; Ersgård 2006; Lucas 2001; Kristiansen 1998)

c) **Report readers do not understand how to use reports** (as discussed by e.g. Aitchison 2010; Harlan 2010; Seymour 2010)

NUMBER OF REPORTS   VALUE OF REPORTS

5

Hypothesis

- **If** there are problems with report readers' understandings of contract archaeology reports hindering efficent usage

then

a more nuanced understanding of the report genre will support future use and management of contract archaeology reports.

6

Aim of the dissertation research

To, through an exploration of **contextual factors** shaping reporting in the intersection of **academic**, **government-professional** and **market values**, provide more

informed grounds for

readings and interpretations of

contract archaeology reports

in order to support

report use and

management.

7

Dissertation = four sub-studies

1) An idea analysis of dimensions in descriptions of grey literature problems in archaeology. (Planned for 2015.)

2) An analysis of information policy in legislation and guidelines regulating contract archaeology reporting. (Ongoing)

3) A bibliographic analysis of the frames of reference report writers bring into reporting. (Finished September 2014.)

4) A focus group study of professional archaeologists' and county board professionals' interpretations of the reporting and report auditing work tasks. (Planned for 2016.)

8

- The dissertation research will be finished by fall 2017.

- The dissertation research is a part of the research project *Archaeological Information in the Digital Society (ARKDIS)* http://arkdis-project.blogspot.se/ and the Center for Digital Heritage http://www.york.ac.uk/digital-heritage/

- Welcome to contact me with questions or comments about the research project! lisa.borjesson@abm.uu.se

ARCHAEOLOGICAL INFORMATION
IN THE DIGITAL SOCIETY
ARKDIS RESEARCH PROJECT
UPPSALA UNIVERSITY | GOTHENBURG UNIVERSITY
LINNAEUS UNIVERSITY | LUND UNIVERSITY

UPPSALA
UNIVERSITET

9

References

Aitchison, Kenneth 2009. "Standards and Guidance in Archaeological Archiving: The Work of the Archaeological Archives Forum and the Institute for Archaeologists." *The Grey Journal* 5 (2): 67–71.

Aitchison, Kenneth 2010. "Grey Literature, Academic Engagement, and Preservation by Understanding." *Archaeologies* 6 (2): 289–300. doi:10.1007/s11759-010-9145-5.

Bahn, Paul G. 2012. *Archaeology: A Very Short Introduction*. Oxford: Oxford University Press.

Darvill, Timothy 2009. "Grey literature", *The Concise Oxford Dictionary of Archaeology.* Oxford: Oxford University Press.

Ersgård, Lars 2006. "Det uppdragsarkeologiska alternativet." In *Det förflutna är att räkna med*, edited by Lars Ersgård, 10–25. Stockholm: Riksantikvarieämbetet.

10

Hardman, Catherine 2009. "The Online Access to the Index of Archaeological Investivations (OASIS) Project" *The Grey Journal* 5 (2): 76-82.

Harlan, Mark E. 2009. "Archaeology's Grey Literature and Professional Review" *The Grey Journal* 5 (2): 83–88.

Heritage Conservation Act (1988:950), revised up to SFS 2104:649

Kristiansen, Kristian 1998. "Gamla gränser och nya frontlinjer." In *Arkeologiska horisonter*, edited by Håkan Karlsson, 176–99. Stockholm: B. Östlings bokförl. Symposion.

Lucas, Gavin 2001. *Critical Approaches to Fieldwork: Contemporary and Historical Archaeological Practice*. London: Routledge.

McGowan, Angela 2009. "Managing the Vulnerabilities of Archaeological Grey Literature". *The Grey Journal* 5 (2): 72-75.

Riksantikvarieämbetet, 2012 *Vägledning för tillämpning av Kulturminneslagen Uppdragsarkeologi (2 kap, 10-13 §§)*, Stockholm: Riksantikvarieämbetet.

Seymour, Deni J. 2010. "In the Trenches Around the Ivory Tower: Introduction to Black-and-White Issues About the Grey Literature." *Archaeologies: Journal of the World Archaeological Congress* 6 (2): 226–32.

Stock, Christiane and Sampité Bernard 2009. "DAPHNE, Full Text Repositories and Other Resources" *The Grey Journal* 5 (2): 97-105.

11

GreyNet Community Management Committee

LinkedIn: 564 members   Twitter: 561 followers   Netvibes: 75 widgets

GreyNet — Grey Literature Network Service

Marcus Vaska
Librarian
Calgary CA

Joachim Schöpfel
Scientist
Lille FR

Hélène Prost
LIS Professional
Nancy FR
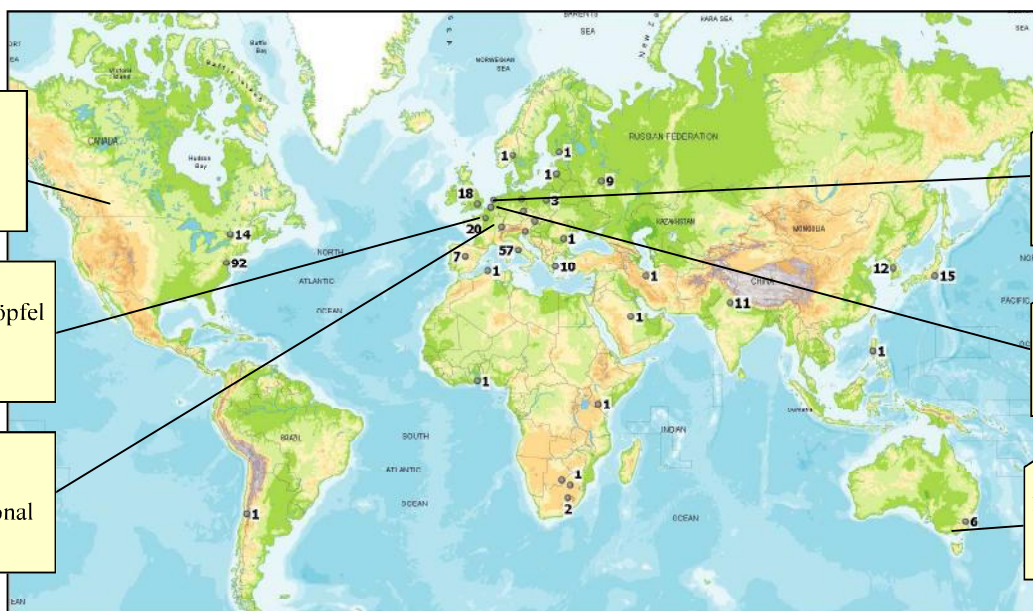
Dominic Farace
Founder GreyNet
Amsterdam NL

Cees de Blaaij
Librarian
Middelburg NL

Amanda Lawrence
Research Manager
Melbourne AU

Objectives:

Increase impact and outreach in social networks
Raise awareness among scientists, scholars and professionals
Motivate them to follow and join the GreyNet community

Twitter: https://   .com/GreyLitNet
LinkedIn: https://   linkedin.com/groups/GreyNet-3718857
Netvibes: http://www.netvibes.com/greynetinternational
Web: http://www.greynet.org
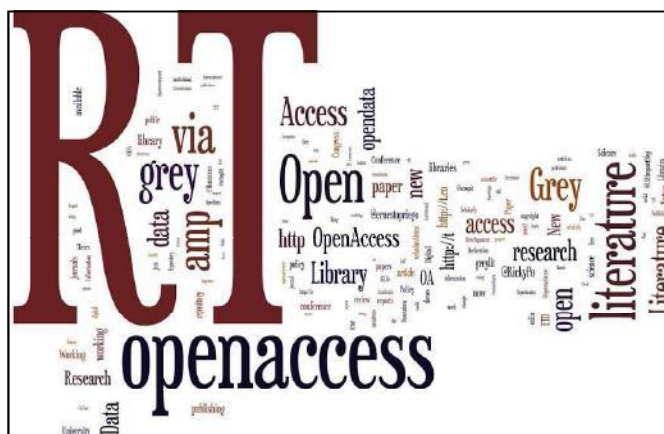
Feedback:

Twitter and *LinkedIn* welcome.
Interested in *open access* and grey literature *collections and sources.*
More information wanted on *open data* and *GreyNet proceedings.*
Moving to *Facebook*?

69

ETD4OA, Electronic Theses and Dissertation for Open Access

Joachim Schöpfel, Helene Prost, Marjorie Piotrowski, University of Lille, France;
Eberhard R. Hilf, Thomas Severiens, Paul Grabbe, Institute for Scientific Networking;
University Oldenburg, Germany

The poster presents a proposal for the European Horizon 2020 Framework Programme for Research and Technological Development. H2020 supports the achievement and functioning of the European Research Area in which researchers, as well as scientific knowledge and technology circulate freely, by strengthening cooperation both between the Union and the Member States, and among the Member States, in particular through the application of a coherent set of rules. Also, a set of activities aims at optimizing the use of national facilities by integrating them into networks and opening their doors to all European researchers.

The project "Electronic Theses and Dissertations for Open Access" (ETD4OA) will support the coordination of European infrastructures and open access (OA) policies in the field of electronic theses and dissertations. Together with stakeholders and OA initiatives, it will address barriers and access restrictions, and it will take actions (active communication, recommendations, advice) to promote and develop input, openness and impact of ETDs in existing open repositories and portals.

The objectives of the ETD4OA project are:

1. to produce and disseminate reliable and consistent empirical data on the processing of electronic theses and dissertations (ETDs) in a wide set of European countries with their different infrastructures and legal environments, in particular to address the problems and variations of access restrictions;
2. to contribute to the coordination of open access policies regarding ETDs in the European research era and to build a sustainable European community focusing on the open access to research theses;
3. to contribute to the understanding and development of the acceptance, uptake and usage of open access ETD infrastructures, by the research communities, by industry and business and by the Internet community;
4. to develop practical and helpful guidelines and recommendations in order to raise awareness about the challenge, and to support cooperation with developing regions and countries, in order to increase the part of ETDs which are widely and fast disseminated in institutional repositories without any restrictions as the standard solution.

The European infrastructure for this already exists, with the interconnection of institutional repositories and other open archives and concurrently registered which can be searched by the DART-Europe portal. The ETD4OA project has not the ambition to develop a new infrastructure or prototype. Instead, today, the problem of open access to ETD lies upstream in local contexts that facilitate decisions in favor of embargoes or restricted access (on-campus access, intranet). To put it in a simple way, pipes exist, but it lacks the fuel and the pressure.

The input to the central portal varies depending on the country and the research field and the institution. To increase the OA-percentage is a problem on many levels, technical (workflow) as well as legal, administrative (regulatory) and tradition and ethics. The answer proposed by the project is awareness raising, lobbying at the national and European policy level, a thorough communication net both for candidates as well as for the responsible local staff, as well as the production of training material for them.

ETD4OA aims at increasing the use of national facilities, i.e. institutional repositories and OA infrastructures, by integrating them into an emerging network of the people, those responsible and involved for European research theses. Its aims are to exploit synergies between national and Union initiatives by setting up partnerships between relevant policy makers and funding bodies or advisory groups, in order to facilitate the development of global research infrastructures and the cooperation of European infrastructures with their non-European counterparts (in particular US and Russia), ensuring their global interoperability and access.

Some of the research questions the project will address: Which are the (apparently) increasing demands by the authors for keeping the document not open to the public for fear of copyright infringement, e. g. of figures used? Is there increasing pressure of commercial publishers to not allow OA distribution if a book is planned? Are there more recent other ways of posting an ETD in OA which are not seen and thus not counted by the institutional repository network, such as for instance institutional servers, personal websites, gold OA journal publications? Language barriers may lead to publish a thesis as a book to make use of the publisher's editorial office.

The project is coordinated by the GERiiCO laboratory at the University of Lille 3 (France) and the ISN Oldenburg (Germany). Actually more than ten member states take part in the project, represented by public or private institutions, experts and networks. ETD4OA has been submitted within the H2020 project call H2020-INFRASUPP-2014-2, "Support to innovation, human resources, policy and international cooperation", a sub call of the H2020-INFRASUPP-2014-2015 program.

Bionotes

Joachim Schöpfel is Lecturer of Library and Information Sciences at the University of Lille 3 (France), Director of the French Digitization Centre for PhD theses (ANRT) and member of the GERiiCO research laboratory. He was Manager of the INIST (CNRS) scientific library from 1999 to 2008. He teaches Library Marketing, Auditing, Intellectual Property and Information Science. His research interests are scientific information and communication, especially Open Access and Grey Literature. Email: joachim.schopfel@univ-lille3.fr

Hélène Prost is information professional at the Institute of Scientific and Technical Information (CNRS) and associate member of the GERiiCO research laboratory (University of Lille 3). She is interested in empirical library and information sciences and statistical data analysis. She participates in research projects on evaluation of collections, document delivery, usage analysis, grey literature and open access, and she is author of several publications. Email: prost@inist.fr

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

Collection and management of science & technology (S&T) information and the development of its service system	Research and analysis of international and local S&T trend
Development and management of a high-performance research network	Development of a high-performance computing infrastructure and its application technology
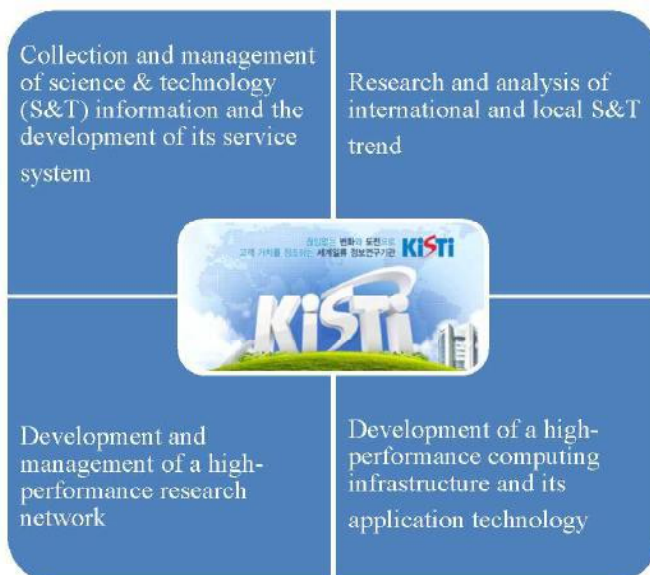
* Management and service of Korean R&D reports

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

Open Access Korea, Phase 1: Five Years On

Hyekyong Hwang, Seonheui Choi, Hyesun Kim;
Korea Institute of Science and Technology Information, KISTI, Korea

Open Access Initiative for the open access to research outputs has been proposed to the world, and in Korea, phase 1 project of Open Access Korea has been conducted with the main strategies of supplying Open Access Korea repositories (hereinafter, "OAK repositories") and Open Access Journal publishing from 2009 to 2013. Phase 2 of OAK (Open Access Korea) has been conducted in 2014 at the National Library of Korea.

This study will introduce the OAK project, phase 1 that has been carried out for the past five years (2009 – 2013). Phase 1 of the project has been conducted by the Korea Institute of Science and Technology Information (KISTI) with the financial support from the Bureau of Library and Museum Policy Planning of the Ministry of Culture, Sports and Tourism. The main activities of OAK are: the development and provision of the OAK repositories; establishment of the XML original of Open Access Journal; Open Access Korea portal services; and dissemination and cooperation activities of Open Access.

The main achievements of OAK for the past five years are the supply of Open Access based OAK institutional repositories to 25 institutions in Korea and the development of OAK repository software as an open source on OAK Portal website, so that all Koreans who wish to operate their own repository may freely download and use the research outputs available on it.

The OAK repositories were developed based on Dspace, which was developed by the Massachusetts Institute of Technology. It supports the Hangul morpheme analyzer apt for circumstances in Korea, provides webpage UI, and supports the automatic linkage with the internal management system thereby enhancing the convenience of management and operation of contents for the repository managers.

Furthermore, through the establishment of the XML original of Open Access Journal, 25 titles of academic society journals were changed to journals in support of the Open Access, and the Korean academic journals are now available to researchers around the world.

The dissemination and cooperation activities of Open Access are as follows:

First of all, we have attempted to legislate for the Open Access Initiatives on the government's research outputs and propose them as legislative bills. For such purposes, judges, lawyers, professors, Creative Commons Korea(CCK) and social activists have drafted legislative bills on public access policy in support of public access of research outputs. With the drafted bills, we have held public hearings to gather the opinions of stakeholders.

Furthermore, we have established expert forums with experts on areas related to Open Access and activists and commenced advocacy activities to support the Open Access Initiatives by hosting voluntary seminars and conferences on related Open Access. In addition, we have established a special group for OAK repository managers for the 25 institutions operating the repositories in order to share know-how amongst them, by holding training sessions and workshops for managers. We have also attempted to raise awareness of Open Access by holding annual OAK(Open Access Korea) conferences.

Through the OAK activities conducted for the past five years, we have achieved the following:

First, we have established the foundation for sharing the knowledge assets of institutions by disseminating the OAK IR suited for Korea.

Second, by support of establishing XML based on JATS 1.10 DTD related to the OA Journal, we have facilitated the open access of Korean research outputs.

Third, through various public relations activities of Open Access, we have contributed to the raise of awareness and activation of Open Access Initiatives in Korea.

Bionotes

Hyekyong Hwang is a senior research librarian of overseas information department in Korea Institute of Science and Technology Information. She is responsible for researching potential Open Access Initiative compatible in Korea Scholarly Communication environment. She earned master of library & information science from Yonsei University in Korea. Her particular interests are in scholarly communications, open access, digital archiving, electronic publishing, and collection development. She is now observing the emergence of a variety of repository frameworks according to open access paradigm. She is a steward of the Open Access Forum in Korea. Email: hkhwang@kisti.re.kr

Hye-Sun Kim is the manager of dept. of NDSL Service at the Korea Institute of Science and Technology Information (KISTI). She has Ph.D in Library and Information Science from Ewha Women's University (2012) and a Master of Art degree in Library and Information Science from Ewha Women's University (1994). Her doctoral dissertation title is 'study on the factors influencing foreign journal subscription in university and college libraries'. Her research interests include: information services, collection development for journals, Open Access, institutional repositories. hskim@kisti.re.kr

75

OAK

Open Access Korea, Phase 1: 5 years on

Dec., 8-9, 2014
Hyekyong Hwang, Seonheui Choi, Hyesun Kim
Korea Institute of Science and Technology Information(KISTI)

Contents

- Project Overview
- Project Outcomes and Activities
  - OAK Institutional Repositories
  - OAK Portal
  - OAK Central
  - Research on Public Access Policy
  - Other Promotional Programs and Activities
- Discussions

2

OAK(Open Access Korea) Project

- **Project Overview**
  - Purpose
    - Facilitating knowledge sharing globally through Open Access
    - Sharing of Domestic Knowledge Information and Developing Advanced Information Environment by Promoting OA Movement in Korea
  - Project period
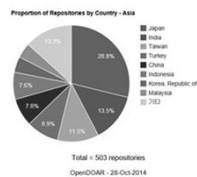   - 5 Years : March 2009-February 2014

3

OAK(Open Access Korea) Project

- Implementing Agency
  - Korea Institute of Science and Technology Information(KISTI)

- Immediate Objectives
  - Promoting open access movement in Korea
    - Green road : Guiding, developing and diffusing of institutional repositories
    - Gold road : Leading scholarly societies and publishers to declaration on open access journals and developing a XML based open access journal publishing system

4

Project Outcomes and Activities

- **OAK Institutional Repositories**
  - Through this project, 23 IRs were developed at selected universities, research institutes, public libraries in Korea and have been being operated since then.

    ❖ Diffusing IRs based on Dspace.
    ❖ Developed OAK IR open source package
    ❖ Special group with OAK repository members
      - Information sharing and cooperation
      - Education and seminars

Proportion of Repositories by Country - Asia

- Japan
- India
- Taiwan
- Turkey
- China
- Indonesia
- Korea, Republic of
- Malaysia
- 기타

Total = 503 repositories
OpenDOAR - 28-Oct-2014

OpenDOAR Repositories by Korea : 7th in Asia
<http://www.opendoar.org/>

5

IR Content Archiving & Dissemination Process

300,000 digital items from 23 IRs
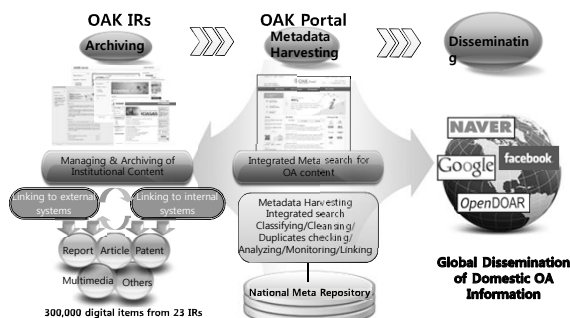
6

Project Outcomes and Activities

• **An OAK Portal**

– which is a search platform for the contents

collected from those 23 IRs, was developed and

has been being operated.

• It holds about 300,000 items of contents(journal articles,

monographs, multi-media etc.) by now.

– Please visit http://oak.go.kr

7

Project Outcomes and Activities

• **OAK Central**

– One of the prominent outputs of the project, is the digital archive for about 25 peer-reviewed, Korean open access journals.

• Constructing full-text according to PMC DTD 3.0

• Developing a work-bench, a semi-automatic tool to convert a PDF file to an XML file

– It contains about 2,979 XML based full-text articles including figures, tables and references

– Exposure to Google, Naver KoreaScience

– See http://central.oak.go.kr

8

Project Outcomes and Activities

• **Research on Public Access Policy**

– KISTI studied on 'public access policy to research outcomes funded by government' for two years from 2010-2011

• Investigated Korean policy and legislative system for the public access to research outcomes

• Investigated public access policies of foreign countries

• Analyzed interest and relationship among stakeholders including researchers, universities, learned societies, and DB companies.

• Prepared a draft bill for public access in Korea

9

Project Outcomes and Activities

- **Other promotional programs and activities**
  - Videos available for explaining about OAK and related activities (See video at http://goo.gl/7oFld)
  - Related educational program has been launched
  - Has hosted both domestic and international meetings and conference under the theme of OA every year
  - An expert committee for OAK has been organized and held the several meetings during the project
    - 20~30 members from government, institutes, universities and DB companies interested in open access.
    - Workshops or seminars to discuss topics on open access.

10

Discussions

- **Gold OA**
  - Understanding of OA awareness is still early stage in KOREA
  - Structurally difficult promoting OA, since it has a lot of stakeholders; authors, publishers, libraries, funders, aggregators, information providers
  - Private DB companies having business of information services are strongly against OA movement
  - Participating in SCOAP3 on behalf of Korean scientific libraries in order to improve accessibility to HEP journals, but we are still develop cost reduction scenarios for SCOAP3 in KOREA.

11

Discussions

- **Green OA**
  - In spite of national and systematical supports, the number of OA IRs was about 23 institutions
    - There has been very slow growth of OA IRs
    - Copyright issues/plagiarism risks/internet security
      - Conflicts with the publishers/societies after sharing in the IR
  - Institutional Written Mandates or Guidelines/Policies are NOT yet established for operations of IRs
    - Several IRs have policies of internal use only
  - Most of IRs are operating by the libraries
    - Cooperation between related internal functional departments, such as the computer team, and research office are needed

12

Big data availability and full exploitation: A 'wide screen' overview on potential opportunities and current criticalities

Luisa De Biagi, Massimiliano Saccone, and Roberto Puccinelli
National Research Council, CNR Central Library, Italy

In recent years a significant effort has been put in fostering the sharing of research data as a mean for increasing the efficacy and efficiency of research fundings. This has been favored by several concurrent and mutually influencing factors, which can be mainly classified as social, political and technological. The collaborative approach, which is common in many scientific fields, has now become a widespread paradigm in many human activities, thanks to the internet technologies, which favor the "social" approach to problem solving. As a consequence, national governments and supernational organizations are devising laws, rules, guidelines and funding instruments in order to provide a framework for an effective sharing of research data (open formats, open licenses, etc.). On the technological side, the availability of broadband connections and the emerging paradigms of big data and cloud computing lay the physical foundations for research data openness.

Nevertheless, there are still some issues, which prevent the full exploitation of the open data potential. Taken for granted the openness of formats and licenses, guarantees are needed as regards data and metadata quality and long term availability and readability. All this implies costs and an accurate planning. More subtle questions concern: which data should be kept and made available (should we preserve only observation data or also experiment data?). On the other hand "big data" availability in open formats provides reuse opportunities to small organizations, which could not afford to produce them through experiments/observations, but does not solve the problems related to the large computing power needed to process them.

The present paper aims at discussing all the above points and to provide a picture of the current state of the debate in Italy, in general, and at CNR, in particular, with references also to the international context.

Bionotes

Luisa De Biagi received her Laurea Degree in Literature and Philosophy at 'La Sapienza' Univ. of Rome (Art history and Cultural heritage). With a Specialization in 'Archivist-Palaeographer' (Vatican School of Palaeography, Diplomatics and Archivistics at the Vatican Secret Archive) as well as a Specialization Degree in Archivistics, Palaeography and Diplomatics (Archivio di Stato, Rome) and a Degree from the Vatican School of Library Sciences. De Biagi further holds a Master in 'Business Publishing' (LUISS Management – Rome). She's been working for the SIGLE Network (System for Information on Grey Literature in Europe) since 2002, since 2010 she is responsible for the Italian National Referring Centre of Grey Literature at CNR Central Library 'G. Marconi' as representative of the European Network and Open archive OpenGrey. She's taken part in 3 editions of the Annual International Congress on Grey Literature (GL5, Amsterdam, GL13 in Washington D.C., and GL14 in Rome). She's is also a member of the CNR Working Group for Cedefop-Refernet Project (Consortium for Professional Education and Training coordinated by ISFOL), the Committee for Legal Deposit Acquisition at CNR Central Library, and a member of the European Association of Health Information and Libraries (EAHIL). She's also responsible for the Library Functional Units 'Education and Training' and 'Cultural Activities Management', organizing didactics laboratories for students, professional training courses and teaching in professional trainings for librarians, students and users. Email: luisa.debiagi@cnr.it

Massimiliano Saccone graduated in Letters and specialized in Library sciences at the "Sapienza" University of Rome. He works as librarian at the Central Library of CNR. He is the responsible for Legal deposit and Open Access activities. He has participated actively in several national and international projects on Information and Knowledge Management (Digital Preservation Europe – DPE, Italian Network of National Bibliography Number – NBN, etc.). His main interests are in scholarly communications, open access, digital preservation, metadata quality control, information system interoperability. Email: massimiliano.saccone@cnr.it

Bionotes *CONTINUED*

Roberto Puccinelli is currently head of Section I at CNR's "Information System Office" and he's been working for CNR since 2001. He has previously worked in the private sector as system and network engineer. As adjunct professor, he has held courses for the First University of Rome "La Sapienza" ("Operating Systems II") and for the Third University of Rome ("Programming and Computing Laboratory"). He graduated in Electronic Engineering at the University of Rome "La Sapienza" and holds a master cum laude in Enterprise Engineering from the University of Rome "Tor Vergata". In the past he has worked in several research projects in the field of Grid technologies both at the national and international level (executive manager of Work Package 11 within the DataGrid project – V Framework Programme, et al.). He's currently involved in the design and development of CNR's information system. In particular, he coordinates projects for the development of application systems and is responsible for the design and implementation of CNR's data warehouse. He is also responsible for CNR's Local Registration Authority management. He's currently involved in projects regarding the design and development of research product open archives and persistent identifier registers/resolvers. He is author of several articles in the fields of Grid technologies, Autonomic Computing, Software Engineering, Open Archives and Persistent Identifiers. Email: roberto.puccinelli@cnr.it

GL 16, Library of Congress,
Washington D.C., December 8–9 2014

**Big data availability and full exploitation.
A 'wide screen' overview on
potential opportunities and current criticalities**

by Luisa De Biagi*, Roberto Puccinelli**,
Massimiliano Saccone*

CNR – National Research Council, Rome (IT)
*CNR – Central Library 'G. Marconi', Rome
** CNR – ICT Office. Network and Informative Systems, Rome

Some preliminary questions

▸ *Too much multid   l matter   the running needing:*

1. A total comprehension of potential opportunities and criticalities connected to big data availability and exploitation.
2. A punctual analysis/monitoring of the research communities and systems involved with their contribution to the production, management, use and re-use of big data. That's possible considering social-economic, technical-technological, cultural, regulatory and policy aspects which actively influence their life cycle and reuse.
3. So, it might be traced an approach and a multidisciplinary-multidimensional study method giving us the opportunity to identify:

1. Stakeholders and contexts
2. Approaches and modalities
3. Topics e objects
4. Instruments

Some preliminary questions/2

As a start, this ideal approach demands a shared language among the different stakeholders and the reference environments of scientific research → it's necessary to identify some fundamental and frequently recurring key-concepts in data research production, management systems and flow.

Thus, it's necessary to find shared or convergent definitions on these terms, at least. Here are some focus *keywords* examples, listed starting from general concepts – which aren't always extensively shared or accepted – to 'focus keywords' expressing more specific concepts.

Scientific communication, Scientific information, Open Science, Open access, Interoperability, access to Research Data, availability, reusability, research data, data set, data collection, e-infrastructure, data infrastructure, data management plan, data policy, big data, curation e/o Preservation, brokerage, data/publications linking, etc.

Legal and ethical issues mainly concern:
➢ intellectual property (industrial Property) considerations
➢ ownership
➢ freedom of Information considerations
➢ privacy laws
➢ data protection laws
➢ human rights considerations

Big Data at Close range

Some goals to pursue	Great challenges
‣ Reducing stakeholder fragmentation ‣ Identifying stakeholder values & inter-relationships ‣ Identifying gaps, tensions & good practice solutions ‣ Identifying and promoting overarching good practice data management/policy solutions ‣ Producing policy guidelines	‣ Stakeholder values & inter-relationships ‣ Infrastructure, technical & technology ‣ Legal and ethical issues ‣ Institutional and policy issues

---

Open Access to Big Data Challenges.
Stakeholder values

‣ **High level policy makers** focus at the very general level and argue for Open Access in terms of very broad social and economic benefits as well as seeing it as a development that will improve science

‣ **Funders** are increasingly motivated to ensure that the allocation of publically funded research yields good value for money. For Research Funders, data is considered a public good, and it must be possible to find and search it.

‣ **Industry** views research data as being potentially any type of data. The viewpoint is that research data should be open on a case-by-case basis, particularly for research that is partially funded by industry. While the former three may be open, the latter as a rule should not be opened except on a case-by-case decision by the owning industry partner. There was reference to the view that data is different from publications and should not be open by default, as the research necessarily is also partially funded by the industry, who is dependent on the commercialisation of the output.

‣ **Stakeholders from within the infrastructural, libraries, repositories**, see value in Open Access to Data as a way of improving the means by which data is made more accessible, and they are motivated to meet the needs   of Open Access within their business cases and service provision

‣ **Publishers** are adapting to the open publishing environment and are developing new types of business models to facilitate that. Here the question of where the cost for Open Access publishing will rest is still undecided

---

Open Access to Big Data Challenges
Infrastructure, technical & technology Issues

In Italy, the Decree No. 91 dated August 8, 2013, stated that the publication of results from **researches with public funds accounting for at least 50% should be stored electronically in** free-access and free-of-charge **repositories**. At the same time, the **Conference of Italian University Rectors (CRUI)** and a number of **public research bodies** have signed a Position statement for the creation of repositories allowing free access to researchers and inciting the same researchers to share their own results, making them available.

Recently CINECA has created **PICO @ Cineca: a new platform for Data Analytics Applications,** a new computing system infrastructure now available at Cineca. PICO is named after Giovanni Pico della Mirandola, Modena, humanist and philosopher of the 1400s, renowned for his prodigious memory. In fact its main goal is responding to the ever-growing demand of services and capacities (storage, management, compute, and visualization) essential to face the challenge of "big data" problems. The system supports data intensive computing cases and features peculiar hardware characteristics (e.g. SSD storage, a type of computer storage media that is made from silicon microchips), software tools, and the high-through put techniques needed by data-oriented projects.

‣ OpenAIRE, a project started in 2009 aimed to: *an infrastructure providing a repository and access to publications under the Seventh Framework Program; develo   virtual place for archiving scientific publications; the creation of an helpdesk network to support researchers →* **OpenAIREplus** (2011), providing a link between **scientific publications** uploaded in the infrastructure **and related information**, including **scientific data.**
Open Data do represent strategic topic, since are free **sharing of research data** for **scientific progress,** but there could be also privacy, security and intellectual property rights **implications.**

‣ With regard to this, the European project **Recode (RECommendations for Open Access to Research Data in Europe),** aimed at **analyzing benefits and issues of open access to data** has been implemented; Italy was represented in this project by CNR Institute of Atmospheric Pollution Research (IIA). "   *le it is easy to define open data, it is not similarly easy to decide how to implement open access to data"* (Stefano Nativi – CNR–IIA).

Open Access to Big Data Challenges
Institutional and policy issues.
Data Acquisition and processing 1/2

Instrument types	Technology examples
API/SPI made available by data provider	Native API/SPI from data provider
ETL * procedures (which may leverage API/SPI calls) * Extract, transform and load	Big Data enabled ETL tools (Pentaho Kettle, Talend ETL, etc.), Hadoop Map Reduce and HDFS and related software like Sqoop, Hive, Pig, Mahout, etc..
Web scraping (capturing data from web pages)	cURL, wget -> page acquisition Apache Tika, Apache POI -> text extraction
Live data stream which may be read continuously (suitable for high-frequency event processing)	Apache Flume, Fluentd, Logstash

Data Acquisition and processing 2/2

‣ **Some issues**:
  ◦ data format heterogeneity,
  ◦ data quality,
  ◦ need for agreements with data providers
‣ **Some lessons learned**:
  ◦ don't ask the provider to adhere to a particular standard unless you are in the position of doing so;
  ◦ alternatively, accept any format and map it to your own;
  ◦ if possible, make available to the provider a tool for metadata mapping, otherwise do the job by yourself.

In real life data may be provided in any possible format and through any possible channel and you have to live with that

Data quality 1/2

‣ **Some data quality aspects**:
  ◦ Completeness: data about entities are complete
  ◦ Consistency: data are not contradictory
  ◦ Accuracy: data are correct
  ◦ Deduplication: entities and related data are represented once and only once
  ◦ Integrity: references between entities are not "broken"

Quality is a process, not a (set of) tool(s)

Data quality 2/2

▸ **Data quality tasks**:
- ◦ Data profiling
- ◦ Data standardization, matching and cleansing
- ◦ Data enrichment
- ◦ Data governance and stewardship

▸ **Data quality tools:**
- ◦ Open source and proprietary products now also integrated with big data technologies (es. Talend, SAP BusinessObjects, etc.)
- ◦ Open source tools developed in the context of funded projects (e.g.: OpenAIRE)

Data storage and access

▸ **Storage platforms**
- ◦ **Distributed file systems**: Hadoop file system
- ◦ **NoSQL data bases**:
  - · Key value stores (ex. FoundationDB, RAMCloud, Voldemort)
  - · Document oriented DBs (ex.: couchDB, Elasticsearch, mongoDB, OrientDB)
  - · Column oriented DBs (ex.: Cassandra, Hbase)
  - · Graph DBs (ex.: Neo4J, Openlink Virtuoso)

▸ **Access requirements and tools**
- ◦ **Discoverability/Harvestability/Searchability → integrate search engines (Solr, Elasticsearch), expose standard interfaces(OAI-PMH, Opensearch)**
- ◦ **Linkability → use Linked Open Data formats**
- **Long Term Preservation → create OAIS compliant repositories**

Technical sustainability and security issues

Sustainability is favored by:
- ◦ use of persistent identifiers;
- ◦ compliance to OAIS reference model.

Security issues:
- ◦ user authentication, authorization, Accounting;
- ◦ privacy and data protection.

Storage platforms and access portals/applications normally have their AAA tools but …

Security is a process, not a (set of) product(s)

Think Tanks, Twitter, and Grey Literature

Danielle Aloia
New York Academy of Medicine, NYAM, United States

This poster aims to track the use of twitter among think tanks and the dissemination of grey literature. Making connections between the future of the online world and knowledge communication to drive change and share resources.

Bionotes

Danielle Aloia received her MSLS from Catholic University of America, in Washington, DC, in 2005 while working on the AgeLine Database at AARP. She has over 20 years of experience in a variety of library settings, including academic, non-profit and museum. She has been involved with collecting, evaluating, and cataloging grey literature beginning in 2006 at AARP and then at the United States. Dept. of Transportation. For the past 3 years she has been managing the Grey Literature Report in Public Health, produced by the New York Academy of Medicine. Email: daloia@nyam.org

Think Tanks, Twitter, and Grey Literature

Sixteenth International
Conference on Grey Literature

Grey Literature Lobby:
Engines and Requesters for Change

December 8-9, 2014

Library of Congress
Washington D.C., USA

Danielle Aloia
NYAM
1216 Fifth Ave.
New York, NY
10029

The New York Academy of Medicine

Grey Literature Report
Fill the gaps in your public health research

Who's on Twitter?

- Think tanks
- Researchers
- Academics
- Journalists
- Individuals

The New York Academy of Medicine

Grey Literature Report
Fill the gaps in your public health research

What do they use it for?

- Engage with community
- Communicate scientific evidence
- Direct communication between researchers/journalists
- Share interests

The New York Academy of Medicine

Grey Literature Report
Fill the gaps in your public health research

Twitter as a measurement tool

- Twitter is being scanned and analyzed to evaluate the effectiveness, reach, and impact of think tanks
- Don't be afraid to retweet:
  Repeating a tweet can really save your news if you had poorly timed it the just time

The New York Academy of Medicine | Grey Literature Report *Fill the gaps in your public health research*

What does the future hold?

- Using Twitter as a dissemination tool is an effective way to disseminate grey literature
- More research needs to be conducted as to the value of Twitter on the dissemination of information

The New York Academy of Medicine | Grey Literature Report *Fill the gaps in your public health research*

Social Media Contact:

- Danielle.aloia.wordpress.com
- @daloia
- Nyamcenterforhistory.com
- @NYAMHistory

The New York Academy of Medicine | Grey Literature Report *Fill the gaps in your public health research*

Capturing from the start: Managing grey literature in a brand new university and enriching the institutional repository

Mohamed Ba-Essa and J.K. Vijayakumar,
King Abdullah University of Science and Technology, KAUST, Saudi Arabia

Institutional repositories are popular solutions in campuses in capturing, preserving and sharing university's intellectual outputs and advance scholarly communication activities. This is an additional channel of dissemination for published materials such as peer reviewed articles, but institutional repositories are the only dissemination mechanism for the grey literature, which bring uniqueness to library collection.

This poster will showcase how a brand new library successfully established an institutional repository, and integrated the grey research outputs like ETDs, reports from the beginning.

We will be discussing, the early consideration of the grey literature, success stories, difficulties, establishing different workflows, plagiarism checks, faculty approvals, stakeholder's involvement, establishing policy and service level agreements. The future challenges associated with grey literature, the increasing demand on different material types such as audio, video and research datasets. The skill sets required in describing gray literature materials, expertise in preserving datasets, research data expertise and continued access are going to be the challenging faces for any repository programs in near future. There are examples of shared grey literature materials in the repository, not only downloaded heavily, but many received citations.

URL: http://repository.kaust.edu.sa/kaust/

Capturing from the start:
managing grey literature in a brand
new research University

Mohamed Ba-essa
J. K. Vijayakumar

Outline

➢ Importance of Research Repository
➢ Importance of Grey Literature
➢ KAUST approach :Importance of policies
➢ The impact
➢ Next Steps

Why KAUST Established
Research Repository ?

➢ Persistent access to university intellectual assets, including grey
  literature (e.g., technical reports, conference papers, theses etc.) and
  research data, in order to preserve and share scientific knowledge
  created at KAUST.
➢ Showcase the intellectual output of KAUST research, development of
  international research networks and collaborations, support graduate
  student and post-doc recruitment.
➢ Expanding impact of KAUST research, which contributes to increased
  awareness of and growing prestige for our new and ambitious
  university.

Why Grey Literature?

Grey literature is also academic literature, but not formally published.

Grey literature is special with more detailed scientific and technical information of any published literature.

Grey literature usually made available before the formal publications.

Research repository is available to preserve and share this valuable information.
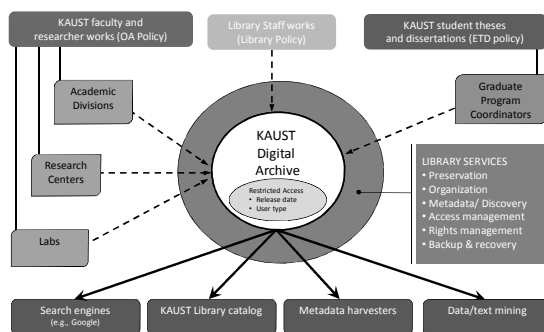
SO, WHY NOT?

KAUST Approach

Policies Built-in

KAUST Dissertation and Thesis Archiving Policy

As a condition of matriculation, KAUST requires that doctoral students submit an electronic copy of their dissertation to the KAUST Library for inclusion in the KAUST digital archive. Similarly, masters students whose program requires a thesis must deposit it in the KAUST digital archive.

KAUST makes no claim of ownership of student dissertations or theses; however, the university retains a non-exclusive license to make copies of dissertations or theses as needed for the academic or archival purposes of the institution. This includes providing open access to the work on the Internet. If necessary to protect legitimate proprietary interests (such as patent rights), students may opt to delay temporarily the public display of their dissertation or thesis.

| *Open Access Policy for Scholarly Research Articles* | Effective Date: June 30, 2014 | *University Library Conference Guidelines: Presentations and papers/abstracts delivered at conferences* | Effective Date: December 2013 |

What is the impact?

Next Steps

➢ Research Data management plan
➢ Supplemental files
➢ Special Collections (biological image collections, GIS images, etc.)
➢ Conference Proceedings

Marine Planning and Service Platform (MAPS)
An Advanced Research Engine for Grey Literature in Marine Science

Sara Goggi, Monica Monachini, Francesca Frontini,
Roberto Bartolini, Gabriella Pardelli, ILC-CNR, Italy;
Maurizio De Mattei, Franco Bustaffa, DP2000, La Spezia, Italy and
Giuseppe Manzella, ETTsolutions, Genova, Italy

The MAPS (Marine Planning and Service Platform) project is a development of the **Marine** project (Ricerca Industriale e Sviluppo Sperimentale Regione Liguria 2007-2013) aiming at building a computer platform for supporting Operative Oceanography in its activities.

One of the main objective of the project is to develop a repository that should gather, classify and structure marine scientific literature and data thus guaranteeing their accessibility to researchers and institutions by means of standard protocols.

Community and Requirements. Operative Oceanography is the branch of marine research which deals with the development of integrated systems for examining and modeling the ocean monitoring and forecast.
Experts need access to real-time data on the state of the sea such as forecasts on temperatures, streams, tides and the relevant scientific literature. This finds application in many areas, ranging from civilian and military safety to protection of off-shore and coastal infrastructures.

The metadata. The set of metadata associated with marine data is defined in the CDI (Common Data Index) documented standard. They encode: the types of sizes which have been measured; the measurement tools the platform which has been employed; the geographic area where measures have been taken; the environmental matrix; the descriptive documentation.
As concerns the scientific documentation, at the current stage of the CDI standard, a document is shaped around the following metadata: *Title, Authors, Version, ISBN/DOI, Topic, Date of publication, Body/Institution, Abstract.*

The search engine. The query system (which is actually under development) has been designed for operating with structured data – the metadata – and raw data – the associated technical and scientific documentation.
Full-text technologies are often unsuccessful when applied to this type of queries since they assume the presence of specific keywords in the text; in order to fix this problem, the MAPS project suggests to use different semantic technologies for retrieving the text and data and thus getting much more complying results.

In the Poster we will present the scenario of the Operative Oceanography together with the technologies used to develop an advanced search engine which aims at providing rapid and efficient access to a Digital Library of oceanographic data. The case-study is also highlighting how the retrieval of grey literature from this specific marine community could be reproduced for similar communities as well, thus revealing the great impact that the processing, re-use as well as application of grey data have on societal needs/problems and their answers.

Essential Bibliography (in progress)

Bustaffa F., De Mattei M.(2013*). MAPS. Raccolta delle esigenze.* . PROGRAMMA OPERATIVO REGIONALE POR-FESR (2007-2013), Asse 1 Innovazione e Competitività, Bando DLTM Azione 1.2.2 "Ricerca industriale e sviluppo sperimentale a favore delle imprese del Distretto Ligure per le Tecnologie Marine (DLTM) anno 2012. CONCESSIONE DI AGEVOLAZIONE POS. N° 19. CUP G45C13000940007 MAPS. Deliverable D1.2,Versione 1.0.

Frontini F., Bartolini R., Monachini M., Pardelli G. MAPS. (2014). Stato dell'Arte dei motori semantici . PROGRAMMA OPERATIVO REGIONALE POR-FESR (2007-2013), Asse 1 Innovazione e Competitività, Bando DLTM Azione 1.2.2 "Ricerca industriale e sviluppo sperimentale a favore delle imprese del Distretto Ligure per le Tecnologie Marine (DLTM) anno 2012. CONCESSIONE DI AGEVOLAZIONE POS. N° 19. CUP G45C13000940007 MAPS. Deliverable D1.1,Versione 1.0.

Bionotes

Sara Goggi is a technologist at CNR-ILC. She started working at ILC in 1996 working on the EC project LE-PAROLE for creating the Italian reference corpus; afterwards she began dealing with the management of several European projects and nowadays she is involved with organisational and managerial activities mainly concerning international relationships and dissemination as well as organization of events (e.g. LREC conference series). Currently one of her permanent activities is the editorial work for the international ISI Journal Language Resources and Evaluation, being its Assistant Editor. Since many years (from 2004) she also carries on research on terminology: this field of research starting intersecting with work on Grey Literature in 2011 at GL13. Email: sara.goggi@ilc.cnr.it

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

Using the PressForward Plugin to
Discover and Manage Grey Literature

Stephanie Westcott, Lisa Rhody, Joan Fragaszy Troyano, and Kim Nguyen
The Roy Rosenzweig Center for History and New Media;
George Mason University, United States

The PressForward Initiative at the Roy Rosenzweig Center for History and New Media (RRCHNM) has developed a methodology and a technology to surface and aggregate grey literature on the open web. Over the past two years we have produced our own open source WordPress plugin to facilitate aggregating, curating, and disseminating scholarly content. Simultaneously, we experiment with multiple processes for surfacing, selecting, and circulating openly published work and grey literature outside formal traditional publication models.

This poster profiles the plugin and highlights the workflows and content of the PressForward prototype publication Digital Humanities Now, offering examples of how organizations can use the PressForward plugin to suit their unique editorial and community needs. Visualizing and documenting intellectual and technical schemas for the plugin with diagrams and onsite demonstrations, this poster exposes the philosophies and methodologies behind aggregating and curating scholarly work manifested in grey literature. Illustrating multiple workflows, layouts, and interfaces, this poster presents the scalable, replicable, and adaptable potential of the PressForward plugin for niche scholarly communities eager to tailor their own hubs for communication and collaboration.

PressForward encourages adaptation of the plugin by others and considers training and educating others about its potential part of its core mission. Exposing the production of our publications through documentation, workflow diagrams, and guides for getting started, this poster offers conference attendees valuable information about an open-source tool that can change how they find, manage, and share grey literature in their field. Through this poster and the conversations it inspires, we hope to encourage viewers to consider how the PressForward model might improve the scholarly communication and collaboration of their own communities of interest on the open web.

GreyGuide

Guide to Good Practices and Resources in Grey Literature

GreyGuide offers information professionals, practitioners, and students common ground for good practice in grey literature

Welcome to the GreyGuide, a repository of good practices and resources in grey literature. The GreyGuide seeks to capture proposed as well as published practices dealing with the supply and demand sides of grey literature. This is a collaborative project involving GreyNet International and ISTI-CNR. The launch of the GreyGuide Repository took place in December 2013 at the Fifteenth International Conference on Grey Literature. Since then, the acquisition of both proposed and published good practices are underway. The GreyGuide is currently still in a developmental stage and is influenced by changes that have taken place in GreyNet's new infrastructure commencing in January 2014. GreyNet is currently in the process of migrating web-based content to the GreyGuide and in effect, the GreyGuide will come to serve as GreyNet's web access portal.

Repository and Web Access Portal

2013
- **Published Good Practices in Grey Literature**
- **Proposed Good Practices in Grey Literature**

2014
- **GL16   Conference Proposals**
- **BIO   Who is in Grey Literature**
- **GSI   GreySource Index**

2015
- **GFS   GreyForum Series**
- **IDGL   International Directory of Organizations in Grey Literature**
- **PUB   GreyNet Inhouse Publications**

Istituto di Scienza e Tecnologie
dell'Informazione "A. Faedo"

GreyNet
Grey Literature Network Service

A Community Driven Open Resource Project in Grey Literature

Pisa Declaration
on
Policy Development for Grey Literature Resources

Introduction

A wealth of knowledge and information is produced by organizations, governments and industry, covering a wide range of subject areas and professional fields, not controlled by commercial publishing. These publications, data and other materials known as grey literature, are an essential resource in scholarly communication, research, and policy making for business, industry, professional practice, and civil society. Grey literature is recognized as a key source of evidence, argument, innovation, and understanding in many disciplines including science, engineering, health, social sciences, education, the arts and humanities.

Grey literature document types in print or electronic formats include among others: research and technical reports, briefings and reviews, evaluations, working papers, conference papers, theses, and multimedia content, representing an important and valuable part of research and information.
In order to realize the benefits of research and information for scholarship, government, civil society, education and the economy, We, the signatories to this declaration, call for increased recognition of grey literature's role and value by governments, academics and all stakeholders, particularly its importance for open access to research, open science, innovation, evidence-based policy, and knowledge transfer.

May 16, 2014

To achieve the full benefits of grey literature for local, national and global communities we call for and encourage the following:

Organizational
1. Greater commitment to open access by governments and organizations.
2. Greater cooperation and coordination among organizations engaged in the production, use, collection and management of grey literature.
3. The use of persistent identifiers and open metadata standards for grey literature.

Research/Educational
4. New forms of recognition and reward for quality grey literature materials by governments, universities and other institutions.
5. Improved standards in the production and bibliographic control of grey literature.
6. Development and implementation of interoperable standards in the management of grey literature.
7. Development of good practice guides for the production, dissemination, and evaluation of grey literature.

Legal
8. Changes to legal deposit and copyright law that enhance the capacities of libraries, collecting services and educational institutions and programs to collect and provide access to grey literature, particularly non-commercial public interest materials.
9. Addressing legal obstacles to the dissemination of grey literature.
10. Further strides in licensing grey content for both commercial and non-commercial purposes.

Financial/Sustainable
11. Identifying available funding for research involving grey literature.
12. Increased support for collection development and long term preservation of grey literature.
13. Increased investment in infrastructure and new technologies for accessing and using print and digital grey literature.

Technical
14. Strategies to tackle link rot and enhance the stability and accessibility of online content.
15. Systems for linking data and other non-textual content to their grey literature publications together with interoperability standards for sharing grey literature.

GreyGuide

To endorse Pisa Declaration go here:

http:// greyguide.isti.cnr.it

GreyNet

Grey Literature Network Service

Contact info@greynet.org

Driving policy development for grey literature – the Pisa Declaration

Herbert Gruttemeier, Chair GreyNet Resource Policy Committee,
Inist-CNRS, France

Current global trends towards open access to research results and open science and innovation provide opportunities to reinforce the role of Grey Literature as an essential resource in scholarly communication. Grey resources are likely to be exemplary in achieving the promises of openness, if supported in different ways, given new forms of recognition and reward, made more widely accessible in a collaborative effort and increasingly linked to various types of scholarly content.

To intensify its promotion of grey resources in this new context, GreyNet has established its Resource Policy Committee, concerned with general issues of strategy, collaboration and policy development. One of the committee's first activities was to prepare a statement of intent, the Pisa Declaration, presented at a seminar in early 2014 hosted by CNR in Pisa and including a comprehensive list of action points addressed to various stakeholders, in the GL communities and beyond. The amended final version of the Pisa Declaration has since been available for signing and commenting on the GreyGuide Repository.

The presentation raises the issue of outreach and impact of the Declaration, recalling some of its key ideas and objectives and, and addressing next potential steps.

Providing access to research data on Africa

Laverne Page
African and Middle Eastern Division; Library of Congress, United States

Description of the collection: The African Section Pamphlet Collection consists of a unique and broad range of ephemeral materials which generally do not meet the scope of the Library of Congress collections guidelines but still may be of considerable scholarly interest for African Studies researchers. These materials are non-commercial, non-book publications in the form of pamphlets, leaflets, broadsides, posters, photographic documentation, policy and technical reports, conference proceedings, data and statistical documents, serials, electoral campaign materials, etc. Some of the material was produced and disseminated by national, regional, and international bodies, and therefore included reports and policy papers of committees and commissions from the Organization of African Unity, the African Union and other African regional organizations. The United Nations and its specialized agencies are also well represented. Included also are documents produced by governmental, academic, business and industry sectors. A wide range of special interest group material is included including political, human rights, refugee, women's, environmental, labor, religious, and youth groups. The collection encompasses both political concerns and a wide variety of topical social, economic, and ideological issues.

The African Section staff has been saving material since 1960. This was the year that the Section was established and a year during which 17 African nations gained independence. Items from the 1960's, printed and distributed in limited quantities, add a rare and unique value to this collection. The collection covers 50 countries and eight regions and is housed in a secure area in approximately 700 archival boxes, arranged geographically, and subdivided by broad subjects. It is a work in progress and is gradually being brought under bibliographic control. Finding aides are being developed for selected parts of this collection.

The African Section Pamphlet Collection supports the research interests of African studies scholars and researchers and can only be accessed in the African and Middle Eastern Division Reading Room.

Bionote

Mattye Laverne Page is an Area Specialist in the African & Middle Eastern Division at the Library of Congress. She has an MA in African Studies from Howard University in Washington, DC and an MLS in Library Science from Columbia in the City of New York. For the Southern Africa region, she is responsible for reference services and collection development, outreach and partnerships with US and overseas libraries, and bibliographic projects. Email: mpag@loc.gov

The recent Italian regulations about the open-access availability of publicly funded research publications, and the documentation landscape in astrophysics

Monica Marra

Istituto Nazionale di Astrofisica, INAF; Osservatorio Astronomico di Bologna, Italy

In October 2013 Italy enacted a law containing the first national regulations about the open-access availability of publicly-funded research results (publications). The impact of these new regulations with the specific situation of that open-access pioneering discipline which is astrophysics, has been considered. Under a strictly technical point of view, in the light of the new dispositions the open nature of a part of the astrophysical scholarly literature which has been made available online free to the reader during the last twenty years, might be questionable. In astrophysics, most of the journal articles are published by a very small number of scientific societies and organizations. The copyright policies of these major publishing bodies have been collected and analyzed, with regard to the main requirements of the Italian law about open access. Most of these policies are sufficiently liberal for an entirely compliant open access to be provided and scientists would benefit from knowing these details more extensively. Some possible ways to make astrophysicists' scholarly dissemination entirely compliant with the Italian law requirements are considered.

Data: Is it Grey, Maligned or Malignant?

Julia Gelfand and Daniel Tsang
University of California, Irvine, United States

Cancers, growths, past events, social issues, conditions, and trends are each proverbially described as on a spectrum from maligned to malignant and scientists, physicians, journalists, commentators, politicians and other specialists offer opinions and commentary on what frames the answer to this question of the title. This paper explores not just the color and tone of data, but attempts to resolve what characterizes whether data is maligned or malignant. Hues of greyness distinguish the perils of failing to share, publish nor make accessible research data and the contemporary consequences to scholarship and open access are critical objectives in today's information arcade. Access to data is determined by those who can afford it, discover and know about it, and can thus manipulate it. Grey literature can take the offensive approach to further the role of data, and promote it to advance the common good, contribute to social responsibility and human actions. Data, while increasingly ubiquitous and abundant is the driver of evidence-based foundations, and the link to academic credibility, communication, discourse, dialogue, and the platform for greater open access. Grey data, possessing some of the attributes of grey literature, difficult to identify, acquire and access, when endangered or threatened, not archived or preserved, requiring methods to organize, sort and stratify, forces nontraditional publishing to pursue data publication to enhance perpetual access and new interpretations for its utility in future learning and research applications. We know that there is a somewhat elevated likelihood that open data policies lead to more widespread knowledge and information sharing, greater self-confidence among information providers and scholars alike, but we know less of whether these patterns have any short or long term benefits or disadvantages for individuals or society and of the factors that moderate and mediate these effects. In the meantime, the new reality is that data is central to the work of science, social sciences and basic human conditions of health and wellbeing and data policies most proceed from a grey containment to this new reality. The argument that as libraries become active publishers by digitizing content, creating new content, supporting researchers by addressing new domains and formats, that other interpretations of grey data and data more generally are increasingly plausible and that further research on the factors moderating and mediating the effects of data management is needed. This paper explores the continuum for data from maligned to malignant and anticipates data approaching the benign stage emphasizing new hues of grey and open access.

Bionote

Julia Gelfand has served as Applied Sciences and Engineering Librarian at the University of California, Irvine Libraries since 1981, where she was recognized as a Distinguished Librarian. She has presented and written extensively on various issues related to grey literature, collection development, scholarly communication, library vendor relationships, emerging technologies, scholarly publishing, social media, and organizational change in academic libraries. She is active in many professional organizations and currently serves on the Board of Directors of the Association of College and Research Libraries (ACRL) Division of the American Library Association (ALA), as Secretary of Section T for Information, Computers & Communication of the American Association for the Advancement of Science (AAAS), and is a member of the Acquisitions & Collection Development Section of the International Federation of Library Associations (IFLA). Email: jgelfand@uci.edu

Daniel C. Tsang has served as Social Sciences Bibliographer and Data Librarian at the University of California, Irvine Libraries since 1986 and was named Distinguished Librarian. He is a well published author in mainstream and alternative media and invited speaker concentrating on anti-surveillance and social movements. Recognized for his scholarship and contributions to the documentation of small and alternative presses, and for his distinguished service to data librarianship, he was awarded the William F. Flanigan Award for Distinguished Service to ICPSR in 2013. He has graduate degrees from the University of Michigan. Email: dtsang@uci.edu

Open data, grey literature and disciplinary differences – Perspectives from a Dutch data archive

Marnix van Berchum
DANS, Data Archiving and Networked Services, The Netherlands

Since 2005 Data Archiving and Networked Services (DANS) promotes sustained access to digital research data. DANS offers several services to support this, including the online archiving system EASY, the Dutch Dataverse Network and the portal NARCIS. In this paper these services will be presented, including the differences we encounter between the disciplines using these services. Within the disciplines served by DANS - mostly belonging to the Social Sciences and Humanities - differences can be discerned regarding the 'openness' of the data. As case study the archeology datasets in EASY will be discussed. Many of them contain reports of archeological excavations done in the Netherlands. Should we consider these as research data or grey literature? And, should we open up the access to these datasets? And if not, why not? These, and other questions, will be addressed, providing a view on how DANS currently deals with open data, grey literature and disciplinary differences.

Bionote

Marnix van Berchum is Head of Data Services at DANS, responsible for the coordination of the services DANS offers (including EASY, the narcis.nl portal, and the Dutch Dataverse Network). After his studies in Musicology at Utrecht University, he started working at Utrecht University Library. In the years 2010-2012 he also worked at SURF. At both employers he was mostly involved with projects related to Open Access and innovations in scholarly communications (e.g. OpenAIRE). Marnix combines his work at DANS with a PhD trajectory at Utrecht University, in which he wants to apply the concepts and methods of network theory on the dissemination of music in the 16th century. Marnix is Associate Director of the CMME Project. Email: marnix.van.berchum@dans.knaw.nl

Data Archiving and Networked Services

Open data, grey literature, and disciplinary differences *perspectives from a Dutch data archive*

Marnix van Berchum
head of Data Services

Washington D.C., 9 December 2014

DANS is an institute of KNAW and NWO

Driven by data

What is DANS?

Institute of Dutch Academy and Research Funding Organisation (KNAW & NWO) since 2005

First predecessor dates back to 1964 (Steinmetz Foundation), Historical Data Archive 1989

Mission: promote and provide permanent access to digital research information

EXPERIENCE 50 YEARS NETHERLANDS

DANS

Our services (i)

DANS
Dutch Dataverse Network

Dutch Dataverse Network
storing data during the research

EASY
Electronic Archiving System

NARCIS
Gateway to scholarly information in the Netherlands

DANS

Our services (ii)

Data Management Plan
for scientific research

How does it work?

Support for Data Management

Data Seal
of Approval

Certification of digital repositories

Projects

persid

Persistent Identifier
URN:NBN resolver

Training & Consultancy

DANS

EASY (i)

- ~ 28k datasets
- certified LTP archive
- Social Sciences
  and Humanities

www.easy.dans.knaw.nl

DANS

EASY (ii)

disciplinary differences

- metadata
- deposit procedures

DANS

**Open if possible,
restricted if necessary!**

DANS

Three (new) categories of access

- Open Access: CC0

- Open Access for Registered Users

- Restricted Access

     + option of embargo period

DANS

Grey literature in EASY

- archaeological reports (excavation results)
- publications belonging to research datasets
- data documentation

GreyNet
Grey Literature Network Service

DANS

Dataset example 1

Dataset example 2

Thank you for your attention

For more information please contact
marnix.van.berchum@dans.knaw.nl

Introduction of "Data Flagging" to utilize numerical data on the Internet - Case of JAEA Fukushima Nuclear Accident Archive

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

In March 2011, the Fukushima Daiichi Nuclear Power Station Accident occurred. Because the information on the accident has been mainly published via the Internet, permanent accessibility to the information, bibliographic control, etc were the critical issues for collecting, preserving and providing Internet information related the Fukushima Accident. We previously reported the development of the Fukushima Archive using the DSpace to solve such issues, and discussed a case of "Internet information is Grey Literature" as a new type of grey literature in Fifteenth International Conference on Grey Literature (GL15).

We have encountered a new challenging issue grey literature. That is utilization of a wide diversity of massive data on the Internet. In many cases, Internet information contains valuable numerical data i.e. raw data, such as radiation monitoring data, plant parameters and physical quantities, and processed data, such as compiled data, evaluated data, experimental data, statistical data and theoretical data and numerical groups in papers.

However, it is difficult to identify the existence of numerical data in the Internet sites because numerical data is not often organized in a systematic manner on the web sites. For example, concerning description of the data, numerical data has multiple items different from literature information such as multiple dates; date created, date modified, date collected, date evaluated etc. It will be very convenient if there is any useful method to identify whether the information contains useful data or not.

We considered such method as to identify numerical data and attempted to introduce "data flagging" system, which was previously used in the International Atomic Energy Agency's International Nuclear Information System (INIS). In case of INIS "data flagging" system, appropriate descriptors were assigned numerical data corresponding to the kind of the numerical data in conjunction with a tag for data flagging. This assigned information helps to identify numerical data which contain the datasheets, table or graphs. In this presentation, we will introduce our "data flagging" system for numerical data on the Internet and how we apply the system to our Fukushima Archive.

Bionotes

Kiyoshi Ikeda works as a librarian at the Central Library of Japan Atomic Energy Agency (JAEA). He also currently works as the chief of International Nuclear Information System (INIS) and Nuclear Information Management Section of JAEA. Since 2011, he has been in charge of selection and classification for INIS Database input data. He joined Japan Atomic Energy Research Institute (JAERI, former JAEA) in 2004, and has been working for acquisition of library materials (domestic/foreign books, journals, technical reports and proceedings). From 2007 to 2010, he had been working for editing work at "JAEA Reports" and "JAEA R&D Review" which are grey literatures. From 2008 to 2013, he has also been a member of editorial committee of the Japan Special Libraries Association (JSLA). He holds a degree in library and information science (M.A.) from the Graduate School of University of Library and Information Science (in Japan).
Email: ikeda.kiyoshi@jaea.go.jp

Bionotes CONTINUED

Mayuki Gonda works as librarian at the Central Library of JAEA (Japan Atomic Energy Agency). He joined JAERI (former JAEA) in 2005, and had been working for management and dissemination of JAEA research results information. Since 2009, he is in charge of selection, classification and indexing for INIS (International Nuclear Information System) Database. He is also a member of editorial committee of the Journal of Information Science and Technology Association (INFOSTA) since 2008. He holds a degree in information science (M.A.) from the Graduate School of Library, Information and Media Studies, University of Tsukuba. Email: gonda.mayuki@jaea.go.jp

Misa Hayakawa works as librarian at the Central Library of Japan Atomic Energy Agency(JAEA). She Joined JAEA in 2010, and has been working for managing metadata of both papers and oral presentations published by JAEA researchers. In addition, she disseminates such information on the Internet via "JAEA Originated Papers Searching System"(JOPSS). She got the master's degree in Library and Information Science at the Graduate School of Library Information and Media Studies, University of Tsukuba in Japan. hayakawa.misa@jaea.go.jp

-

Introduction of "Data Flagging" to utilize numerical data on the Internet
- case of JAEA Fukushima Nuclear Accident Archive -
Kiyoshi Ikeda et al. *Japan Atomic Energy Agency (JAEA), Japan*

The Possibilities of Numerical Data

is

infinite

16th International Conference on Grey Literature, 8-9 December 2014, Washington D.C.

Background

- Tokyo Electric Power Company (TEPCO) Fukushima Daiichi Nuclear Power Station accident ("Fukushima Accident") occurred on March 11th, 2011.
- Since then, the Japanese government, research institutions, TEPCO, etc. disseminate a variety of information.
- We discussed a case of "Internet information can be Grey Literature" as a new type of grey literature in 15th International Conference on Grey Literature (GL15)

Grey Literature

Records/Documents (report, press release), Photos/Movies, Numerical data (radiation monitoring, plant parameters) etc.

2

Collection of Fukushima Accident Information

- We started to provide reference Information related to the Fukushima Accident in April, 2011
- We cover conventional literature for International Nuclear Information System(INIS) Database as much as possible
- Many information are published via Internet but they are difficult to be covered
- In 2013, we started to collect, organize and transmit such information
- "Fukushima Nuclear Accident Archive" was launched on June 23th, 2014

- List of the R&D results related to Fukushima Accident by JAEA staffs

- Bibliographies related to the TMI-2, Chernobylsk-4 reactor accidents, environmental decontamination etc.

- Internet sources related to the Fukushima Accident

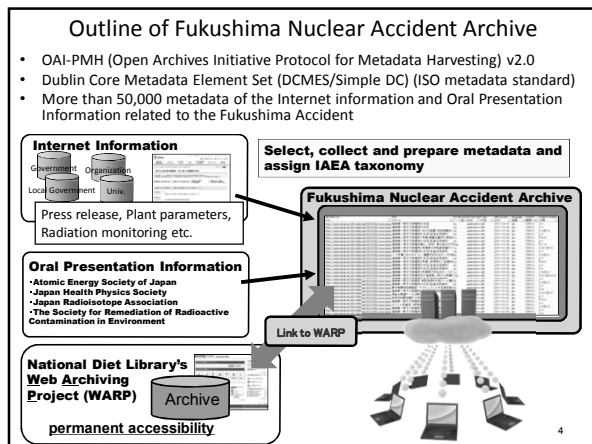**JAEA Library Website**

3

Outline of Fukushima Nuclear Accident Archive

- OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) v2.0
- Dublin Core Metadata Element Set (DCMES/Simple DC) (ISO metadata standard)
- More than 50,000 metadata of the Internet information and Oral Presentation Information related to the Fukushima Accident

Internet Information
Government   Organization
Local Government   Univ.
Press release, Plant parameters, Radiation monitoring etc.

Select, collect and prepare metadata and assign IAEA taxonomy

Fukushima Nuclear Accident Archive

Oral Presentation Information
·Atomic Energy Society of Japan
·Japan Health Physics Society
·Japan Radioisotope Association
·The Society for Remediation of Radioactive Contamination in Environment

Link to WARP

National Diet Library's Web Archiving Project (WARP)
Archive

permanent accessibility

4

Fukushima Nuclear Accident Archive

http://dspace.jaea.go.jp

Numerical Data

- In many cases, Internet information contains valuable numerical data, such as observed data, compiled data, evaluated data, experimental data, statistical data, theoretical data etc.

- In the field of nuclear science and technology, for example, useful numerical data is radiation monitoring data; plant parameters; physical quantities such as densities, melting points, cross sections, spectral shifts; statistical and probability data; or numbers of occurrences of certain events such as failures of pieces of equipment etc.
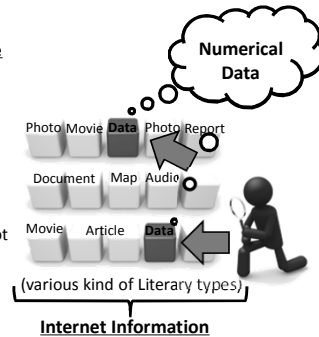
Numerical Data

Issue

Difficult to identify the existence of numerical data in the Internet sites because …

- ◦ Numerical data is not often organized in a systematic manner on the web sites
- ◦ In many cases, there is no information in the metadata which indicates whether it contains numerical data or not

Something is necessary in the metadata to identify whether it contains numerical data or not
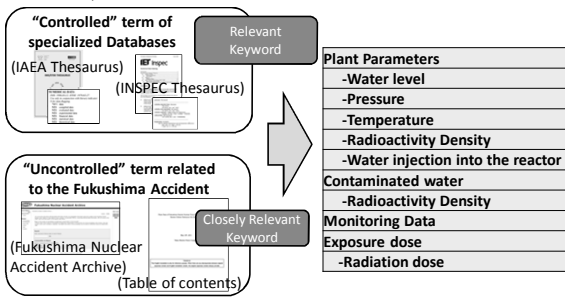
Numerical Data

Photo Movie Data Photo Report

Document Map Audio

Movie Article Data

(various kind of Literary types)

Internet Information

Data Flagging in INIS

- To make a search for Internet information containing numerical data easier, we considered such method as to identify numerical data and attempted to introduce "data flagging" system, which is used in the International Atomic Energy Agency's International Nuclear Information System (INIS).

- In INIS "data flagging", appropriate descriptors are assigned corresponding to the kind of the numerical data in conjunction with a tag for Data Flagging. This assigned information helps to identify numerical data.
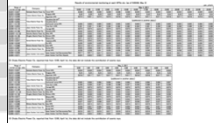
Data Flagging 一 Keyword

- We have chosen the following keywords for "Data Flagging". Table of contents of *Plant Data of Fukushima Daiichi NPS at the time of the Tohoku-Chihou-Taiheiyou-Oki Earthquake*, Fukushima Nuclear Accident Archive, INIS Thesaurus and INSPEC Thesaurus are used as reference

"Controlled" term of specialized Databases

(IAEA Thesaurus)

(INSPEC Thesaurus)

Relevant Keyword

"Uncontrolled" term related to the Fukushima Accident

(Fukushima Nuclear Accident Archive)

(Table of contents)

Closely Relevant Keyword

Plant Parameters
-Water level
-Pressure
-Temperature
-Radioactivity Density
-Water injection into the reactor
Contaminated water
-Radioactivity Density
Monitoring Data
Exposure dose
-Radiation dose

Data Flagging in Fukushima Archive

- We put the following information as "Data Flagging"
  - Put "numerical data" as tag in Type field which indicates the record contains numerical data
  - Add keywords in Subject.chapter field which indicates the kind of the numerical data included

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	Add Tag: "Numerical Data"
date.issued	2011-05-03
language	eng
creator	METI
subject.chapter	Add Keywords: "Monitoring Data" etc.

Summary

- Internet information contains valuable numerical data, such as compiled data, evaluated data, experimental data, statistical data, theoretical data etc.

- We introduce "Data Flagging" for easily searching numerical data in Fukushima Nuclear Accident Archive

- The user can significantly benefit by appropriate indications of data content of Internet information related to the Fukushima Accident by the processes of Data Flagging

From Japan to the world

Thank you very much

Moving towards Open Access to research data in health sciences: a changing landscape for Grey Literature

Paola De Castro, Istituto Superiore di Sanità, Italy

The traditional role of GL in health sciences will be discussed and new perspectives for GL exploitation will be considered against the general trend of open data sharing within the open science movement. In the past, one of the major concerns about GL was its limited circulation and availability. Today this concern is generally overcome thanks to the Internet, but GL authors and issuing organizations face new responsibilities, mainly associated with the scientific value of GL and new sharing opportunities. The traditional borderline between commercial and non-commercial or grey literature, which was very clear in the past century, is now becoming obsolete, in fact the Open Access (OA) movement to scientific publications implies the free (libre) use of different types research output, and therefore it also includes GL, as stated in the numerous declarations in support of Open access, stating from the year 2003.

In the past century, technical reports represented the only type of documents able to host large amounts of data which journal articles, mainly for economic reasons, were not able to print. Today journals are even encouraged to publish enriched publications including datasets and they are looking for the best way to share data so as to be re-used for different purposes. In this regard also GL producers can play a major role in the new information sharing arena, together with journals editors, research funders, authors, and other stakeholders who will benefit from a new sharing attitude. Some examples in health sciences will be provided stressing that today it is very important to create awareness on the advantages of sharing information and data; such awareness may be promoted through the organization of ad hoc initiatives addressed to the different stakeholders, including policy makers, to show how data sharing can contribute to accelerate progress in science, maximize return of investment, develop of cross-disciplinary research, improve a general understanding of research efforts, etc.

Examples of current open access policies, recommendations and regulations issued by research institutes, funding agencies and governments will be also provided to show how the new trends towards open data sharing do include GL as a relevant primary source of research information.

In this regard the direct involvement of the grey community in the international debate on open data becomes fundamental to share positive and negative experiences and contribute to speed up progress towards new forms of information dissemination, definitively cleaning up the grey dust and poor quality often associated with GL.

Bionotes

Paola De Castro is Director of the Publishing Unit of ISS, the National Institute of Health in Italy. Active supporter of Open Access to scientific publications and research data in public health, promoting training and research activity in support of research dissemination and cooperation at different levels. She is involved in different projects to improve the editorial quality of grey literature, as well as its dissemination, mainly through training courses addressed to scientists and other actions addressed to different stakeholders, at national and international level. Coordinator of the EU FP7 Project NECOBELAC. Member of the Italian task force of the FP7 MedOAnet project in support of Open Access policies in Southern Europe. She organizes conferences, workshops and exhibitions on public health issues addressed to both health operators and to the general public, including students. She promotes and carries on projects on the preservation of historical heritage of scientific institutions (including scientific instruments, unpublished documents, pictures, oral memories). Council member of the European Association of Science Editors (EASE), country representative of HIFA2015, member of the European Association of Health Information and Libraries (EAHIL) and AIB. She has published many articles on information communication and research evaluation; she organizes and delivers face to face and virtual courses on scientific communication at national and international level. Served the EU Commission as evaluator of FP7 Projects and acted as peer reviewers for other research projects and journal articles. Email: paola.decastro@iss.it

Changes in the International Nuclear Information System

Dobrica Savić, Nuclear Information Section, NIS; Department of Nuclear Energy,
International Atomic Energy Agency, IAEA, Austria

In 1970, the International Nuclear Information System (INIS) became the IAEA's first nuclear database. Ten years later, online access was offered, and in 1998, it was the first IAEA database on the Internet. The biggest change occurred in 2009, when Internet users were given free and open access to this resource. Within 5 years, the number of users increased ten-fold, reaching well over a million visitors annually, by making the INIS Collection searchable through Google and Google Scholar. Just as the IT revolution changed the environment in which we work, changes that would take place within INIS information services were as yet unforeseen. Change became the only constant in our daily work. As the speed of change increases, the outcome becomes even harder to predict.

Currently, INIS hosts one of the world's largest collections of published information on the peaceful uses of nuclear science and technology, offering online access to a unique collection of 3.7 million bibliographic records, and over a million full texts (non-conventional (grey) literature). Annually, there are over 3 million INIS page-views and more than a million online visitors. This is the result of changes implemented over the years in the management of INIS.

In order to prepare for the future, changes in the nature of grey literature and in the provision of information services must be reviewed, along with the current state of technology and other megatrends.

Bionote

Dobrica Savić is Head of the Nuclear Information Section (NIS) at the International Atomic Energy Agency (IAEA) in Vienna, Austria, which comprises of the International Nuclear Information System (INIS), the IAEA Library, and the IT support group. Mr. Savić 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. He 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 modern information technologies. Email: d.savic@iaea.org

LinkedIn

GreyNet

Grey Literature Network Service

GreyGuide, GreyNet's web access portal and lobby for Grey Literature

Dominic Farace and Jerry Frantzen, GreyNet International, Netherlands
Stefania Biagioni, **Carlo Carlesi and Roberto Ponti,** ISTI-CNR, National Research Council, Italy
Christiane Stock, Inist-CNRS, National Centre of Scientific Research, France

In December 2013, the GreyGuide was formerly launched as an online forum and repository of good practice in grey literature. The project partners then turned to the acquisition of both proposed and published good practices. During this same timeframe, GreyNet– one of the project partners – welcomed far reaching developments in its infrastructure. Three new committees were established in line with its fourfold mission dedicated to research, publication, open access, and education in the field of grey literature.

 In the process of coordinating and facilitating the work of these new committees, it became clear that a multitude of web-based content, which is currently maintained on GreyNet's website and the GL conference site, should be made accessible via a unique web portal. This would allow for browse, search, and retrieval across collections. The GreyGuide was tested for this purpose and it was then decided to select and migrate content to the GreyGuide.

While the web origins of four such collections have been identified

- GreySource Index
- Who's Who in Grey Literature
- International Directory of Organizations in Grey Literature
- GL Conference Proposals

The work of defining the metadata for these collections, their subsequent data entry, and additional cross-linking indicate the work that will be undertaken in the coming months. It is anticipated that just as GL15 provided the occasion for the launch of the GreyGuide Repository, GL16 will demonstrate its enhanced function as a web access portal.

Bionote

Dominic Farace is Head of GreyNet International and Director of TextRelease, an independent information bureau specializing in grey literature and networked information. He holds degrees in sociology from Creighton University (BA) and the University of New Orleans (MA), and has a doctorate in social sciences from the University of Utrecht, Netherlands – where he lives and works since 1976. After six years heading the Department of Documentary Information at the Royal Netherlands Academy of Arts and Sciences (1986-1992), Farace founded GreyNet, Grey Literature Network Service. He has since been responsible for the International Conference Series on Grey Literature (1993-2015). In this capacity, he also serves as Program and Conference Director as well as managing editor of the Conference Proceedings. He is editor of The Grey Journal and provides workshops and training in the field of grey literature. Email: dominic.farace@textrelease.com

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

Christiane Stock is the Head of the Monographs and Grey Literature service at INIST, in charge of the repositories LARA (reports), mémSIC (master's theses in information sciences) and OpenGrey. Member of the Technical Committee for the SIGLE database from 1993 to 2005, she also set up the national agency for ISRN (International Standard Report Number). She is member of the AFNOR expert group who prepared the recommended metadata scheme for French electronic theses (TEF). Email: christiane.stock@inist.fr

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

GreyGuide - GreyNet's Web Access Portal
and Lobby for Grey Literature

Dominic Farace and Jerry Frantzen, GreyNet International

Stefania Biagioni, Carlo Carlesi, Roberto Ponti, ISTI-CNR, Italy

Christiane Stock, Inist-CNRS, France

Istituto di Scienza e Tecnologie
dell'Informazione "A. Faedo" | GreyNet Grey Literature Network Service | **Inist** **cnrs**

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

2013
Repository of Proposed and Published
Good Practices in Grey Literature

2014
GreyNet's Web Access Portal

2

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

Guide to Good Practices and Resources in Grey Literature

GreyNet's Web based Resources and Collections:

• GreySource Index

• WHOIS in Grey Literature

• International Directory of Organizations in Grey Literature

• GL Conference Proposals

• GreyForum Series

• In-house Publications

3

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

GreyNet's New Infrastructure in 2014

Research and Development	Publishing and Communication
Program Committee and Chair ①	Community Management Committee and Chair ②
International Conference Series on Grey Literature	*Serials, Non-Serials, and Social Media*
Open Access	Education and Training
Resource Policy Committee and Chair ③	LIS Committee and Chair ④
Repositories and Web-based Resources	*Library and Information Science*

4

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

Resource Policy Committee - RPC Meeting
11 October 2013 at Inist-CNRS, Nancy, France

The main objective for this meeting was to bring together key stakeholders in the field of grey literature, who are working with GreyNet in the design and development of open access repositories better serving the needs of research communities, educational institutions, and the wider public. Another objective for this meeting was to seek further collaboration among these organizations both on technical as well as promotional levels.
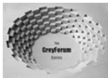
5

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

Grey Literature and Policy Development: The Pisa Declaration
Seminar held on 7 April 2014 at ISTI-CNR in Pisa, Italy

This GreyForum focused on 'Grey Literature and Policy Development' related to open access, open data, open standards, open science, etc. Former SIGLE members, current OpenGrey partners, as well as other stakeholders both within and outside the grey literature community were invited to actively participate.

6

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

Resource Policy Committee - RPC Meeting
16 May 2014 at Inist-CNRS in Nancy, France

Tasks:

• To finalize the text of the Pisa Declaration on Policy Development for
Grey Literature Resources and publish it via the GreyGuide

• To commence lobbying for the endorsement of the Pisa Declaration

• To review metadata for GreyNet's Resources and Collections
that await migration to the GreyGuide Repository

7

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

Pisa Declaration on Policy Development for Grey Literature Resources
"A New Roadmap for Grey Literature"

http://greyguide.isti.cnr.it

• The Document • To endorse • List of Endorsements

8

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

GreyGuide Management Team Meeting
30 June – 2 July 2014 at ISTI-CNR, Pisa, Italy

Tasks:

• To prioritize GreyNet's input in the GreyGuide Repository

• To better engineer the technical and management workflow of the GreyGuide

• To commence with the record input of GL16 Conference Proposals
and to continue with the migration of GreySource entries in the GreyGuide

9

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

GreyGuide Management Team Meeting
6 - 7 October 2014 at ISTI-CNR, Pisa, Italy

Tasks:

• To streamline input in the GreyGuide

• To review and revise lead texts, field names, and help screens
consistent with the inclusion of GreyNet collections

• To directly engage the grey literature community by requesting them
to enter their own record in the BIO Collection "Who is in Grey Literature"

10

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

Now a Year on and the Way Forward

• Some Initial Results and Statistics

• Emphasis on Community Sourcing to populate the GreyGuide

• Exploitation of the GreyGuide portal function
through contact with advanced gateways and discovery services

• Lobbying public awareness of grey literature by way of the GreyGuide

11

Sixteenth International Conference on Grey Literature
Library of Congress, Washington D.C. December 8-9, 2014

Grey**G**uide ● ● ●

Thank You!

For more information visit
http://greyguide.isti.cnr.it

Contact:

Stefania Biagioni, Repository Manager
stefania.biagioni@isti.cnr.it

Dominic Farace, Project Manager
info@greynet.org

12

Grey Lit Tutorials

Rachel Smiley, Sarah Davis, and Debbie Rabina
Pratt Institute, School of Information and Library Science, United States

Last December, at GL15, GreyNet International introduced the education and training working group that was charged with providing instructional materials to grey lit users. We are now pleased to present the beta version of Grey Lit Tutorials.

Grey Lit Tutorials (http://greylittutorials.wordpress.com) is an educational outreach website devoted to explaining and exploring grey literature. Through videos, current news, and event information, Grey Lit Tutorials provides a better understanding of how academics can be best served through this new, hybrid model of publishing. The website runs in tangent to GreyNet International (http://www.greynet.org), and serves as a user-friendly method of information dissemination.

The heading bar on Grey Lit Tutorials provides a quick summary of the important features of the website. "About Grey Lit Tutorials" is the first place for visitors to learn more about the website as a whole and the joint project between GreyNet International and Pratt Institute. The next section, "What is Grey Lit?," explains more about grey literature as a publishing model, and includes an instructional video. "Video Tutorials" is the home of all videos produced by the project team about grey literature, and is updated regularly. A descriptive list of relevant grey literature sites and organizations are listed under "Annotated Collections." Finally, all events, publicity notices, and press releases related to grey literature can be found under the "Events" tab. The user can return to the home page through the banner at the top, where the site's most recent news and posts can be found.

When planning the structure of this grey literature research guide, the LibGuide platform from SpringShare guided the site design (http://springshare.com). Since the use of LibGuides is proprietary, however, the free software platform available from Wordpress was chosen as an alternative (http://wordpress.com). This hosting choice also allows more format flexibility, personal customization, and institutional independence for future subject-specialty administrators.

Grey Lit Tutorials provides an educational, user-friendly companion website to GreyNet International, with original content as well as a republished pieces from the GL community. The dynamic, flexible nature of Grey Lit Tutorials will make the website an easily-searchable site for anyone interested in grey literature and will continue to supplement the growing interest in new publishing methods. We encourage users to fill out the form on the Grey Lit Tutorials website and suggest collections or other materials you would like to see included. As always, we look forward to your feedback and participation to help build this resource we can all use.

Bionote

Debbie Rabina is associate professor at Pratt Institute, School of Information and Library Science. Her research and teaching are within the areas of information policy, official information and scholarly communication, and informed by the belief that societies that exercise a policy of open dissemination of and access to information are those that allow their citizenry to achieve personal and societal goals. Debbie is immediate past member of the Depository Library Council, an advisory board to the Public Printer of the United States, heads and World Information Societies and Environments concentration at Pratt and is a member of the Content working group of the Library of Congress, National Digital Stewardship Alliance. She publishes regularly in the LIS field, most recently in Journal of Documentation, JELIS, Advances in Librarianship and The Grey Journal. Email: drabina@pratt.edu

Publishing and Communication: The GreyNet Community Management Committee

Joachim Schöpfel, Chair
Charles de Gaulle University Lille 3
ANRT – GERiiCO, France

The GreyNet Community Management Committee (CMC) currently consists of six members: Joachim Schöpfel Chair (France), Cees de Blaaij and Dominic Farace (The Netherlands), Amanda Lawrence (Australia), Hélène Prost (France), and Marcus Vaska (Canada).

Although several people work and publish in the field of grey literature, few know about GreyNet International and its events, publications and services. In order to increase impact and outreach in social networks, GreyNet International launched a Community Management Committee (CMC) at the end of 2013. The goal of the CMC is to raise awareness among scientists, scholars and professionals, and motivate them to follow and join the GreyNet community. Priority is given to three social media outlets: LinkedIn (social network), Twitter (microblogging) and Netvibes (personalized dashboard publishing platform). Enabling an RSS feed from GreyNet news is also on the agenda.

The paper will introduce the Committee, and present its objectives and priorities. It will also show the results after the first year of community management, in terms of followers of the Twitter account, members of the LinkedIn group, etc.

In order to obtain feedback from the GreyNet community (members, sponsors etc.), a short survey will be launched a couple of months prior to the GL16 conference, focusing on acceptance, satisfaction and suggestions of the CMC; results will be included in the poster and presented to conference participants.

Bionotes

Joachim Schöpfel is Lecturer of Library and Information Sciences at the University of Lille 3 (France), Director of the French Digitization Centre for PhD theses (ANRT) and member of the GERiiCO research laboratory. He was Manager of the INIST (CNRS) scientific library from 1999 to 2008. He teaches Library Marketing, Auditing, Intellectual Property and Information Science. His research interests are scientific information and communication, especially Open Access and Grey Literature. Email: joachim.schopfel@univ-lille3.fr

Cees de Blaaij studied Social and Economic History at the University of Nijmegen and Library Science at the University of Amsterdam. He worked for Ernst & Young, management consultants, and the Institute for Information Law (University Amsterdam) as information professional. At the moment he is working as an academic librarian and coordinator digital services for the Public and Academic Library of Zealand, Netherlands. He took part in several GL conferences. He published several articles on issues concerning copyright in the digital environment and accessibility of grey literature on the Internet. Email: cdbl@zebi.nl

Amanda Lawrence is a researcher with Swinburne University's Institute for Social Research. Ms Lawrence is managing editor of Australian Policy Online and a guest lecturer in information architecture at the Royal Melbourne Institute of Technology (RMIT). She holds a Graduate Diploma in Library and Information management (RMIT) and BA (Hons) Arts (Melbourne). She is qualified for membership of the Australian Library and Information Association (ALIA). Email: alawrence@swin.edu.au

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

Marcus Vaska is a librarian with the Knowledge Resource Service (KRS), University of Calgary, Holy Cross Site, providing research and information support at an Alberta Cancer Care research facility. A firm supporter of embedded librarianship, Marcus engages himself in numerous activities, including instruction and research consultation, with research teams at Holy Cross and beyond. Marcus' current interests focus on collaborative educational techniques aimed at creating greater awareness and bringing grey literature to the forefront in the medical community. Email: mmvaska@ucalgary.ca

Royal Netherlands Academy of Arts and Sciences
Amsterdam, Netherlands, December 1 - 2, 2015

A New Wave of Textual and Non-Textual Grey Literature
Seventeenth International Conference on Grey Literature

Appendices

List of Participating Organizations

African and Middle Eastern Division; Library of Congress	United States
American Psychological Association, APA	United States
Aerospace Corporation, The	United States
Biblioteca Centrale, G. Marconi; CNR	Italy
Canadian Agency for Drugs and Technologies in Health, CADTH	Canada
Centre National de Recherché Scientifique, CNRS	France
Dalhousie University, School of Information Management	Canada
Data Archiving and Networked Services, DANS	Netherlands
Delta Progetti 2000 S.r.l.; COMDATA	Italy
EBSCO Publishing	United States
ETT Solutions Ltd.	Italy
Federal Library Information Network, FEDLINK	United States
Federal Reserve Bank of St. Louis	United States
GERiiCO laboratory	France
Grey Literature Network Service, GreyNet International	Netherlands
Indiana University School of Informatics and Computing, IUPUI	United States
Information International Associates, IIa	United States
Institut de l'Information Scientifique et Technique, Inist	France
Institute of Computational Linguistics, ILC-CNR	Italy
Institute of Information Science and Technologies, ISTI	Italy
Institute for Scientific Networking, ISN Oldenburg	Germany
Istituto Nazionale di Astrofisica, INAF; Osservatorio Astronomico di Bologna	Italy
Istituto Superiore di Sanità, ISS	Italy
Japan Atomic Energy Agency, JAEA	Japan
King Abdullah University of Science and Technology, KAUST	Saudi Arabia
Knowledge Resource Service, KRS	Canada
Korea Institute of Science & Technology Information, KISTI	Korea
Library of Congress, LoC	United States
National Library of Technology, NTK	Czech Republic
National Research Council, CNR	Italy
National Union Catalog of Manuscript Collections, NUCMC	United States
New York Academy of Medicine, NYAM	United States
Nuclear Information Section; International Atomic Energy Agency, NIS-IAEA	Austria
Pratt Institute, School of Information and Library Science	United States
PricewaterhouseCoopers, PwC	Netherlands
Research Institute of Geodesy, Topography and Cartography, VUGTK	Czech Republic
Roy Rosenzweig Center for History and New Media, George Mason University	United States
Slovak Centre of Scientific and Technical Information, CVTISR	Slovakia
Slovak University of Technology in Bratislava, STUBA	Slovakia
Swinburne University of Technology, SWIN	Australia
TextRelease, Program and Conference Bureau	Netherlands
Tisch Library, Tufts University	United States
Université Charles de Gaulle - Lille 3	France
University Hospitals Bristol, BRI; NHS Foundation Trust	United Kingdom
University of Calgary	Canada
University of California, Berkeley School of Information	United States
University of California, Irvine Libraries, UCI	United States
University of Lille 3	France
Uppsala University	Sweden
University of Wisconsin-Milwaukee, School of Information Studies	United States
U.S. Department of Justice, Civil and Criminal Division	United States
U.S. Department of Transportation, DOT	United States

Index to Authors
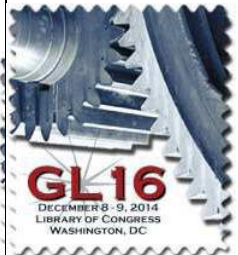
Grey Literature Lobby: Engines and Requesters for Change

The Library of Congress, Washington D.C. USA on 8-9 December 2014

Publication Order Form

SIXTEENTH INTERNATIONAL CONFERENCE ON GREY LITERATURE

Publication(s):	No. of Copies	x	Amount in Euros	Subtotal
GL16 CONFERENCE PROCEEDINGS – Printed Edition ISBN 978-90-77484-23-4 ISSN 1386-2316 *Postage and Handling **excluded***)*		x	109.00 =	€
GL16 CONFERENCE PROCEEDINGS - CD-Rom Edition ISBN 978-90-77484-23-4 ISSN 1386-2316 *Postage and Handling **included***		x	109.00 =	€
GL16 Conference Proceedings – Online Edition ISBN 978-90-77484-23-4 ISSN 2211-7199 ***Password Protected Access***		x	109.00 =	€

POSTAGE AND HANDLING PER PRINTED COPY*)

Holland		x 5.00	
Europe		x 10.00	€
Other		x 15.00	€
		TOTAL EURO =	**€**

Customer Name:	
Organisation:	
Postal Address:	
City/Code/Country:	
E-mail Address:	

Upon receipt of payment the publication(s) will be forwarded to your shipping address with an invoice marked paid.

❑ Direct transfer to TextRelease, Account No. 3135.85.342, Rabobank Amsterdam
BIC: RABONL2U **IBAN:** NL70 RABO 0313 5853 42, with reference to "GL16 Publication Order"

❑ MasterCard/Eurocard   ❑ Visa Card   ❑ American Express

Card No. _____   Expiration Date: _____

Print the name that appears on the credit card, here _____

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

Place: _____   Date: _____

NOTE: CREDIT CARD TRANSACTIONS CAN BE AUTHORIZED BY PHONE, FAX, OR POSTAL SERVICES. EMAIL IS NOT AUTHORIZED.