

# Fourteenth International Conference on Grey Literature

## National Research Council, Rome, Italy 29-30 November 2012

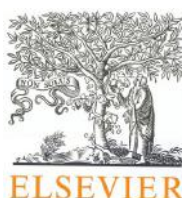
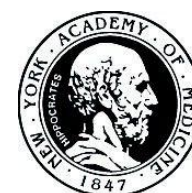
TRACKING INNOVATION THROUGH GREY LITERATURE



# Program Book

ISSN 1385-2308

## Host and Sponsors



GL14 Program and Conference Bureau

**TextRelease**

Javastraat 194-HS, 1095 CP Amsterdam, The Netherlands  
[www.textrelease.com](http://www.textrelease.com) • [conference@textrelease.com](mailto:conference@textrelease.com)  
Tel/Fax +31-20-331.2420



## **CIP**

### **GL14 Program Book**

Fourteenth International Conference on Grey Literature : Tracking Innovation Through Grey Literature, 29-30 November 2012 / compiled by D.J. Farace and J. Frantzen ; GreyNet, Grey Literature Network Service. Amsterdam : TextRelease, November 2012. – 166 p. – Author Index. –

(GL-conference series, ISSN 1385-2308 ; No. 14)

CVTISR, EBSCO, FEDLINK, INIST-CNRS, NTK, and NYAM are corporate authors and associate members of GreyNet International. This program book contains the schedule for the plenary and poster sessions. The titles and abstracts of the papers as well as information on the authors are included. When available, copies of the Powerpoint slides are provided in notepad format.

ISBN: 978-90-77484-19-7

© TextRelease, 2012

## Foreword

# **TRACKING INNOVATION THROUGH GREY LITERATURE**

Innovation is a process manifested in and through grey literature. Both have their origins in knowledge generation and both demonstrate value for government, academics, business and industry through their uses and applications. In a way, innovation and grey literature are two sides of the same coinage. Innovation is the catalyst for positive change and grey literature is the measure of benchmarks in the further process of research and development. Innovation and grey literature share parallel life cycles in which early growth is relatively slow until their use and application become recognized both within and later beyond their community of origin. Expected top-line growth and increased bottom-line results are achieved in part through new technologies, through redeployment and enhancement of existing products and services, which at times are unachieved. Nevertheless, the process shared by innovation and grey literature carries on.

The goal of the Fourteenth International Conference on Grey Literature set out to track the process of innovation by tracing the research life cycle and observing its convergence in the field of grey literature. Thirty-five presentations from authors and researchers from 17 countries worldwide are harvested in this Program Book.

*Dr. Dominic J. Farace*  
Grey Literature Network Service

Amsterdam,  
November 2012

## GL14 Conference Sponsors



CNR, National Research Council, Italy  
Central Library "Guglielmo Marconi"



IRPPS, Italy  
Institute of Research on Population and Social Policies  
National Research Council, CNR



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



FEDLINK, USA  
Federal Library Information Network  
Library of Congress



EBSCO Publishing, USA



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



JST, Japan  
Japan Science and Technology Agency

## GL14 Conference Sponsors



Reed Elsevier, Netherlands



THOMSON REUTERS

Thomson Reuters, USA

**NTK**

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

NTK, Czech Republic  
National Technical Library



NYAM, USA  
New York Academy of Medicine



**Springer**

Springer Verlag, Germany



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



KISTI, Korea  
Korea Institute of Science and Technology Information



*GreyNet*

Grey Literature Network Service



## Table of Contents

	Foreword.....	3
	Conference Sponsors.....	4
	Program and Planning Committee Members.....	8
	Conference Outline .....	9
	Conference Program and Timetable .....	10
<b>Program</b>	Opening Session.....	13
	Session One: Tracing the Research Life Cycle.....	21
	Session Two: Tracking Methods for Grey Literature.....	43
	Poster Session and Sponsor Showcase.....	58
	Session Three: Adapting New Technologies.....	120
	Session Four: Repurposing Grey Literature.....	145
<b>Advertisements</b>	GreyNet LinkedIn.....	6
	J-STAGE – Japan Science and Technology Agency.....	20
	KISTI, Korea Institute of Science and Technology Information.....	36
	EBSCO LISTA Full-Text.....	42
	Refdoc.fr, INIST-CNRS.....	56
	GreyForum Series <sup>NEW</sup> .....	84
	NYAM, The New York Academy of Medicine.....	97
	NTK, National Technical Library, Czech Republic.....	98
	CVTISR, Slovak Centre of Scientific and Technical Information.....	132
	De Gruyter Saur, <i>Grey Literature in Library and Information Studies</i> .....	150
	GL15 Pre-Conference Announcement.....	164
<b>Appendices</b>	List of Participating Organizations .....	163
	Index to Authors.....	165
	GL14 Publication Order Form.....	166

## GL14 Program and Planning Committee

### Planning Committee:



Rita Ciampichetti, Chiara D'Arpa, Raffaellaalle, Adelaide Ranchino, Flavia Cancedda, and Luisa De Biagi  
National Research Council, Italy

### Program Committee:



Stefania Biagioni **Chair**  
Institute of Information Science and Technologies, ISTI  
National Research Council, Italy



Rosa Di Cesare and Daniela Luzi, **Co-Chair**  
Institute of Research on Population and Social Policies, IRPPS  
National Research Council, Italy



Danielle Aloia  
New York Academy of Medicine, USA



Christiane Stock  
Institut de l'Information Scientifique et Technique, INIST  
Centre National de Recherche Scientifique, France



Roberta Shaffer  
Library of Congress, USA



Petra Pejřšová  
National Technical Library, Czech Republic



Joachim Schöpfel  
University of Lille 3, France



Blane Dessy  
Federal Library and Information Center Committee, USA



Dominic Farace  
Grey Literature Network Service, Netherlands





## TRACKING INNOVATION THROUGH GREY LITERATURE

## Conference Outline

<b>DAY 1 Thursday</b>	<b>DAY 2 Friday</b>
<b>Location:</b> Piazzale Aldo Moro 7, Rome, Italy	<b>Location:</b> Piazzale Aldo Moro 7, Rome, Italy
<b>Registration Desk Open</b> 8:00-9:00 Coffee and Tea Service	<b>Registration Desk Open</b> 8:30-9:00 Coffee and Tea Service
<b>Opening Session</b> 9:00-10:30 <i>Welcome Address: Prof. Luigi Nicolais, CNR President</i> <i>Opening Address: Dr. Carlos Morais-Pires, EU Commission</i> <i>Keynote Address: Dr. Jan Brase, TIB Hannover</i> <i>Rejoinder: Dr. Luisa De Biagi and Flavia Cancedda, CNR</i>	<b>Poster Session and Sponsor Showcase</b> 9:00-11:00 <i>Posters are scheduled to be presented in an informal setting adjacent to the main conference hall</i>
<b>Morning Break</b> 10:30-11:00	
<b>Session One</b> 11:00-13:00 'Tracing the Research Life Cycle'	<b>Session Three</b> 11:00-12:30 'Adapting New Technologies'
<b>Lunch</b> 13:00-14:00	<b>Lunch</b> 12:30-13:30
<b>Session Two</b> 14:00-15:30 'Tracking Methods for Grey Literature'	<b>Session Four</b> 13:30-15:00 'Repurposing Grey Literature'
<b>Afternoon Break</b> 15:30 -16:00	<b>Pause</b> 15:00-15:15
<b>Introduction to Posters</b> 16:00-17:00 <i>Short briefings by the Poster Presenters</i>	<b>Closing Session</b> 15:15-16:00 <i>Report by Chairpersons, Conference Handoff, Farewell</i>
<b>Conference Reception</b> 17:00-18:30 <i>Including Special Guest Presentation</i>	<b>CNR Post-Conference Tour</b> 16:00-17:00

## DAY ONE

9:00 – 10:30

## OPENING SESSION

**Chair**, Stefania Biagioni, Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", ISTI-CNR, Italy**Welcome Address** Prof. Luigi Nicolais, President, Consiglio Nazionale delle Ricerche, CNR, Italy**Opening Address** **E-Infrastructure challenges to organise the collaborative European scientific data space** 13  
Dr. Carlos Morais-Pires, European Commission, Brussels, Belgium**Keynote Address** **Paradigm shifts in Information Access - Beyond classical scholarly publication** 14  
Dr. Jan Brase, Head DOI Registration Agency; Technische Informationsbibliothek, Germany**Rejoinder** **Being persistent and identifiable in the digital era: Persistent Identifiers as a strategic tool to valorize grey literature and research products** 15  
Dr.ssa Luisa De Biagi and Flavia Cancedda, CNR Central Library 'G. Marconi', Italy

10.30 -11:00 Morning Break

11:00 – 13:00

## SESSION ONE – TRACING THE RESEARCH LIFE CYCLE

**Chair**, Donatella Castelli, Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", ISTI-CNR, Italy**Customized OAI-ORE and OAI-PMH Exports of Compound Objects for the Fedora Repository** 21  
Alessia Bardi, Sandro La Bruzzo, and Paolo Manghi; Istituto di Scienza e Tecnologie dell'Informazione, CNR, Italy**Grey literature in the digital culture and practices of the new global scholar: the case of molecular biology** 27  
Chérifa Boukacem-Zeghmouri; Département d'Informatique de l'Université Claude Bernard Lyon 1, France**Grey in the Innovation Process** Keith G Jeffery, STFC-Rutherford Appleton Laboratory, United Kingdom and 29  
Anne Asserson, University of Bergen, Norway**Characteristics and use of grey literature in scientific journal articles of Algerian University of Science and Technology teachers and researchers in STM fields** 34  
Lydia Chalabi, Research Center of Scientific and Technical Information, CERIST, Algeria**The Research Life Cycle and Innovation through Grey Literature in Nanotechnology in Korea** 37  
Seon-Hee Lee, Korea Institute of Science and Technology Information, KISTI, Korea

13:00-14:00 Lunch

14:00 – 15:30

## SESSION TWO: TRACKING METHODS FOR GREY LITERATURE

**Chair**, June Crowe, Information International Associates, IIA, USA**What goes up must come down: Publications from developing countries in the Aquatic Common** 43  
Maria Kalentsits and Armand Gribling; Fisheries & Aquaculture Branch Library, FAO - Food and Agriculture Organization of the United Nations, Italy**Data sharing in environmental sciences: A survey of CNR researchers** Daniela Luzi; Institute for Research on 48  
Population and Social Policies, IRPPS-CNR, Roberta Ruggieri, Senato della Repubblica, Stefania Biagioni, Institute of Information Science and Technologies, ISTI-CNR, Elisabetta Schiano, Institute of Marine Sciences, ISMAR-CNR, Italy**Tracking the Influence of Grey Literature in Public Policy Contexts: The Necessity and Benefit of Interdisciplinary Research** Bertrum H. MacDonald, Elizabeth M. De Santo, Kevin Quigley, Suzuette S. Soomai, and Peter G. Wells; 54  
Dalhousie University, Canada**Grey communities: An empirical study on databases and repositories** Hélène Prost; Institute for Scientific and 57  
Technical Information, INIST-CNRS and Joachim Schöpfel, Charles de Gaulle University Lille 3, France

15.30 -16:00 Afternoon Break

16:00 – 17:00

## INTRODUCTION TO CONFERENCE POSTERS

**Chair**, Dominic Farace, Grey Literature Network Service, GreyNet, Netherlands

17:00 – 18:30

## CONFERENCE RECEPTION

Special Presentation by Prof. Emeritus Paul Sturges, United Kingdom

## DAY TWO

9:00 – 11:00

## POSTER SESSION AND SPONSOR SHOWCASE

Conference posters are scheduled to be presented in an informal setting adjacent to the main conference hall

58

*Delegates from Algeria, Australia, Belgium, Canada, Czech Republic, France, Germany, Greece, India, Italy, Japan, Korea, Netherlands, Norway, Slovak Republic, United Kingdom, and United States*

10.30 -11:00 Morning Break

11:00 – 12:30

## SESSION THREE – ADAPTING NEW TECHNOLOGIES

**Chair,** Daniela Luzi, *Institute for Research on Population and Social Policies, IRPPS-CNR, Italy*

**An Environment Supporting the Production of Live Research Objects** *Massimiliano Assante, Leonardo Candela, Donatella Castelli, and Pasquale Pagano; Istituto di Scienza e Tecnologie dell'Informazione – CNR, Italy*

120

**Creating and Assessing a Subject-based Blog for Current Awareness within a Cancer Care Environment** *Yongtao Lin and Marcus Vaska; Health Information Network Calgary, University of Calgary, Canada*

126

**Centralised National Corpus of Electronic Theses and Dissertations**

133

*Július Kravjar and Marta Dušková; Slovak Centre of Scientific and Technical information, CVTISR, Slovakia*

**A funder repository of heterogeneous grey literature material with advanced user interface and presentation features** *Ioanna-Ourlana Stathopoulou, Nikos Houssos, Panagiotis Stathopoulos, Despina Hardouveli, Alexandra Roubani, Ioanna Sarantopoulou, Alexandros Soumplis, Chrysostomos Nanakos; National Documentation Centre; National Hellenic Research Foundation, Greece*

139

12.30 -13:30 Lunch

13:30 – 15:00

## SESSION FOUR – REPURPOSING GREY LITERATURE

**Chair,** Bertrum H. MacDonald, *Dalhousie University, Canada*

**Working for an open e-publishing service to improve grey literature editorial quality**

145

*Rosa Di Cesare, Marianna Nobile; Institute for Research on Population and Social Policies, IRPPS Silvia Giannini; Institute of Information Science and Technology, ISTI, Italy*

**Teaching and learning across national frontiers to improve dissemination of scientific research output: Is there a borderline between grey and white literature? Lessons learned from the NECOBELAC experience in Europe and Latin America** *Paola De Castro, Istituto Superiore di Sanità, ISS, Italy*

151

**Grey literature in Australian education** *Gerald White, Julian Thomas, Paul Weldon, Amanda Lawrence and Helen Galatis; Australian Council for Educational Research, Australia*

157

**Research Life Cycle: Exploring Credibility of Metrics and Value in a New Era of eScholarship that Supports Grey Literature** *Julia Gelfand, University of California, Irvine (UCI) and Anthony Lin, Irvine Valley College, USA*

162

15:00-15:15 Pause

15.15 – 16:00

## CLOSING SESSION – REPORTS CHAIRPERSONS, CONFERENCE HANDOFF, FAREWELL

**Chairs,** Stefania Biagioni, *ISTI-CNR, Italy* and Dominic Farace, *GreyNet International, Netherlands*

16:00-17:00

## POST CONFERENCE TOUR - CNR CENTRAL LIBRARY

Guided by Flavia Cancedda and Luisa De Biagi, CNR Biblioteca Centrale "G. Marconi", Italy

## POSTER SESSION AND SPONSOR SHOWCASE

<b>Use pattern of archives on the history of Mysore</b>	<b>58</b>
<i>N. Chowdappa, L. Usha Devi, C.P. Ramasesh, BMS College of Engineering, India</i>	
<b>Data Analytics: The next big thing in information</b>	<b>63</b>
<i>June Crowe and J.R. Candlish, Information International Associates, Inc.</i>	
<b>GreyNet's Enhanced Publications Project: Tracking and Backtracking Data</b>	<b>68</b>
<i>Dominic Farace and Jerry Frantzen, GreyNet International, Netherlands; Christiane Stock, INIST-CNRS, France; Paula Witkamp, Data Archiving and Networked Services, DANS, Netherlands</i>	
<b>Collection of Conference Proceedings and Improving Access to the Full Text of Proceedings</b>	<b>73</b>
<i>Misa Hayakawa, Shun Nagaya, Mayuki Gonda, Takeyasu Fukazawa, Minoru Yonezawa, and Keizo Itabashi, Japan Atomic Energy Agency, JAEA, Japan</i>	
<b>Envia: Environmental Information Discovery and Access</b>	<b>78</b>
<i>Johanna Kieniewicz, Elizabeth Newbold, and Alan Sudlow, Science Technology and Medicine, The British Library, United Kingdom</i>	
<b>Grey Literature, E-Repositories, and Evaluation of Academic &amp; Research Institutes: The case study of BPI e-repository</b>	<b>79</b>
<i>Maria V. Kitsiou and Vasileios Souvlidis, Benaki Phytopathological Institute, Greece</i>	
<b>Innovation, language, and the web</b>	<b>85</b>
<i>Claudia Marzi, Institute for Computational Linguistics, "Antonio Zampolli", National Research Council, CNR, Italy</i>	
<b>Open Grey for Natural Language Processing: A ride on the network</b>	<b>89</b>
<i>Gabriella Pardelli, Sara Goggi, Manuela Sassi, Istituto di Linguistica Computazionale, "Antonio Zampolli", ILC-CNR, Italy</i>	
<b>IRPPS Editoria Elettronica: An electronic publishing web portal based on Open Journal Systems (OJS)</b>	<b>95</b>
<i>Marianna Nobile and Fabrizio Pecoraro, Institute for Research on Population and Social Policies, IRPPS, Italy</i>	
<b>Grey literature partnership network in the Czech Republic</b>	<b>99</b>
<i>Petra Pejřová and Hana Vyčítalová, National Technical Library, Czech Republic</i>	
<b>Building the evidence base to support food and feed safety risk assessments in Europe</b>	<b>103</b>
<i>J. Richardson, L. Congiu, C. Morganti, P. Pirro, E. Aissa, S. Noorani, D. Lefebvre, and D. Verloo, European Food Safety Authority, Italy</i>	
<b>Tools and resources supporting the cultural tourism</b>	<b>104</b>
<i>Eva Sassolini, Alessandra Cinini, Stefano Sbrulli, and Eugenio Picchi, Istituto di Linguistica Computazionale, ILC, Italy</i>	
<b>Semantic networks for improved access to biomedical databases</b>	<b>108</b>
<i>Eva Sassolini, Sebastiana Cucurullo, and Eugenio Picchi, Istituto di Linguistica Computazionale, ILC, Italy</i>	
<b>Five Ways to Add Value to PhD Theses and Two Ways to do it Better</b>	<b>114</b>
<i>Joachim Schöpfel, Charles de Gaulle University Lille 3; UFR IDIST &amp; ANRT, France</i>	

## Opening Address

### **E-Infrastructure challenges to organise the collaborative European scientific data space**

**Dr. Carlos Morais-Pires**

Head Sector Scientific Data Infrastructures  
European Commission, Belgium



**Carlos Morais Pires** (European Commission) is Head of the Sector of the 'Scientific Data Infrastructure' in the unit dealing with e-Infrastructures in the EU 7th Framework Programme (Capacities Programme). He has been with the European Commission (Information Society DG) since 1998, dealing with the development of a large number of R&D projects on Information and Communication Technologies and has been actively involved in bringing output from research into policy and regulatory activities. He holds a Ph.D. from Telecom Italia Lab/University of Lisbon, in Electrical Engineering in the field of Telecommunication and Broadcasting Technologies. Before joining the EC, he was with the department of Computer Science of ISCTE University in Lisbon leading the group on computer networks, operating systems and multimedia. Email: [carlos.morais-pires@ec.europa.eu](mailto:carlos.morais-pires@ec.europa.eu)

## Keynote Address

### Paradigm shifts in Information Access – Beyond classical scholarly publication

**Dr. Jan Brase,**

Head DOI Registration Agency  
Technische Informationsbibliothek, Germany



#### **Abstract**

Scientific information today can be found in various forms and content types: grey literature, data sets, videos, 3D models, etc. Nevertheless for hundreds of years the scholarly publications system has been based upon the classical article publication and citation system. The information community must now rise to the challenges of the so-called 4th paradigm by developing solutions together with the scientific community to make any kind of scientific information available, citable, sharable, linkable and usable by the use of persistent identifiers.

## Bionote

---

**Jan Brase** has a degree in Mathematics and a PhD in Computer science. His research background is metadata, ontologies and digital libraries. Since 2005 he is head of the DOI-registration agency for research data at the German National Library of Science and Technology (TIB). He is furthermore Managing Agent of DataCite, an international consortium with 16 members from 12 countries <http://www.datacite.org>. DataCite was founded in December 2009 and has set itself the goal of making online access to all kinds of information for scientists easier by promoting the acceptance of research data and grey literature as individual, citable scientific objects. DataCite has already assigned over 1.3 million DOI names to data sets and grey literature. Jan is Chair of the International DOI foundation (IDF), Vice-President of the International Council for Scientific and Technical Information (ICSTI) and Co-Chair of the recently established CODATA Data Citation task group. He is author of several articles and conference papers on the citation of data sets and the new challenges for libraries in dealing with such non-textual information objects.

Email: [Jan.Brase@tib.uni-hannover.de](mailto:Jan.Brase@tib.uni-hannover.de)

## Rejoinder

### Being persistent and identifiable in the digital era: Persistent Identifiers as a strategic tool to valorize grey literature and research products

**Dr.ssa Luisa De Biagi**, CNR Central Library 'G. Marconi'

**Dr. Flavia Cancedda**, Italian Centre ISSN c/o CNR Central Library 'G. Marconi', Italy

Long-term availability of digital objects, especially e.g. online dissertations causes new challenges. This rewards particularly libraries with a legal mandate for collecting, archiving and indexing national cultural heritage. Several aspects must be considered in order to guarantee long-term access to the digital resources, e.g. the preservation of the bit stream of digital objects, realizing migration, and emulation strategies. But a fundamental requirement to achieve long-term availability is the use of Persistent Identifiers (PIs) which assures unequivocal resource identification and reliable addressing. In fact, although a persistent identifier does not in itself ensure persistent access, it is a necessary step along the way to ensuring a "shelf mark" for distributed internet resources, aiding in long-term preservation, and a mechanism by which resources can be related to one another in a stable way. In this way, resources can enrich each other for the benefit of users who wish to use these connections. Identifiers hold the key which will enable information resources to be related to one another in a stable and meaningful way. This is what the basis of the "semantic web" is all about, or what has also been described as "Web 3.0". In fact, Google is still not able to find and connect together resources which are directly (semantically) related to one another. These relationships have to be encoded by people or advanced machines, and at their basis will have to be persistent and unambiguous resource identifiers, for any document type and any target, both people and organizations. Hence, this paper aims possibly to trace a state of the art on the PIs implementation, with special regard to the Italian situation and web 2.0/3.0 electronic resources, both in institutional/public administration and industry, entrepreneurship, 3rd sector etc. Following, an 'empirical' study on the use and demand of ISSN code as a strategic tool for valorizing Italian research products and outcomes, in relation to the recent disposals on the research quality evaluation (Gelmini Law 240/2010).

## Bionotes

**Luisa De Biagi** received her degree in Literature and Philosophy at 'La Sapienza' Univ. of Rome 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). Degree from the Vatican School of Library Sciences. De Biagi further holds a Master in 'Business Publishing' (LUISS Management – Rome). She was a member of the SIGLE Working Group and Italian Grey Literature Data-Base Working Group from 2002 and a Member of the CNR Working Group for Cedefop-Refernet Project (Consortium for Professional Education and Training coordinated by ISFOL). Member of the Committee for Legal Deposit Acquisition at CNR Central Library, and European Association of Health Information and Libraries (EAHIL). Since 2010 is Responsible for the Italian National Referring Centre of Grey Literature at CNR Central Library 'G. Marconi' (OpenSIGLE/OpenGrey Network Project) and 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](mailto:luisa.debiagi@cnr.it)

**Flavia Cancedda** graduated in Literature at the University of Rome, specialized in Librarianship (Manuscript preservation) in the same University and obtained a Master's degree in Library Direction and Management at the Cattolica University (Milan). She worked as a librarian in two different Italian universities and in the Central National Library of Florence. She is currently working in the Central Library of the Italian National Research Council, in Rome, where collaborated with the Grey Literature National Centre from 2002 to 2006. Since 2006 she is responsible for the Italian ISSN Centre, located in the same Library, and is member of the 'Documentation and Information Committee – Subcommittee 9, Identification and description' – homologous for ISO TC46-SC9 - in the UNI (Italian Standardization Organization). Her areas of interest and publishing activity are related to the book history, historical bibliography, standards for descriptive and subject cataloguing, semantic classifications, standards for identifiers and persistent identifiers. Email: [flavia.cancedda@cnr.it](mailto:flavia.cancedda@cnr.it)

Consiglio Nazionale delle Ricerche

***Being persistent and identifiable in the digital era. Persistent Identifiers as a strategic tool to valorize grey literature and research products***

---

---

---

---

---

---

---

Consiglio Nazionale delle Ricerche

▪ **By Luisa De Biagi\*, Flavia Cancedda\*\***

\* Consiglio Nazionale delle Ricerche (CNR)  
Informative and Documental Systems Office.  
National Research Council Central Library 'G. Marconi'

*National Centre for Grey Literature\**  
-  
*National ISSN Centre\*\**

---

---

---

---

---

---

---

**Planet Digital Libraries:  
key features and benefits**

*For definition DL provide a wide sightseeing of all information contained within a library, no matter its format or support, and serve special communities through the network*

**DL main features:**

- collaborative work and virtual workplaces;
- digital document preservation;
- distributed database management;
- hypertext;
- information filtering and information retrieval;
- instructional modules;
- intellectual property rights;
- multimedia information services;
- FAQ and reference services;
- informative resources discovery and selective dissemination of information

---

---

---

---

---

---

---



## Planet Digital Libraries: main key features and benefits

### Digital libraries benefits and advantages (compared with traditional libraries):

- Constant availability of up-to-date, high quality multimedia resources → easier removal of physical and conceptual barriers;
- Network connectivity and interactive technologies → Rich virtual workplaces empowering social interactions;
- Digital technology → advanced and innovative services, usually hard and expensive for a traditional library

---

---

---

---

---

---

---

---

## DL standards and metadata

- *Metadata is structured information for a better retrieval, use or management of information resources. The term metadata is used differently in different communities.*
- *In libraries metadata is commonly used for formal scheme of resource description, applying to any type of digital object. (MARC 21 - and its rule sets AACR2 - is one of the metadata standard).*
- For the practical purposes, metadata can be classified into three broad categories:
- **Descriptive Metadata:** describe and identify information resource through the use of search tools, as well as provides sufficient context for understanding what has been found. The best-known descriptive metadata standard for libraries is MARC (Machine-Readable Catalog) and Dublin Core
- **Structural metadata:** enhance navigation and presentation of electronic resources
- **Administrative meta-data:** facilitate both short-term and long-term management and processing of *digital collections*

---

---

---

---

---

---

---

---

## DL Persistent Identifiers: main current types and qualities

- **DL need persistent naming** → digital resources must remain the same regardless of where the document is located. Anytime, anywhere (No matter if it changes location or ownership changes)

PI must be:

- **Unique (permanent naming system)**
- Global
- Reliable

→ Unique name is a basic quality for:

- citations
- information retrieval
- links among objects
- managing copyright

---

---

---

---

---

---

---

---

## DL Persistent Identifiers: main current types and qualities

- Uniform Resource Name (URN)
- Internationalized Resource Identifier (IRI)
- Persistent Uniform Resource Locator (PURL)
- Digital Object Identifier (DOI)
- Archival Resource Key (ARK)
- CNRI Handle System (<http://hdl.handle.net>)
- OAI Identifier

---

---

---

---

---

---

---

---

## DL Persistent Identifiers: other kind of construction hypothesis

### *Why not?*

In grey literature:

The old 'ISRN' [dead in 2007 in ISO environment, but alive in old catalogues and repositories as Opengrey] ->> could be transmigrated in new DOI identifiers

In white literature:

The traditional 'ISSN' ->> is *involved* as the core element in the construction of several kinds of Persistent identifiers

In a general scenario of identifying people:

The new ISNI (International Standard Name Identifier) ...

---

---

---

---

---

---

---

---

## DL Persistent Identifiers: ISSN as including element

According to ISO 3297:2007

ISSN and/or ISSN-L identifiers could be incorporated in (*examples*):

- DOI
- OpenURL
- URN
- ...

▪ In other identifiers:

- SICI
- EA N-GS1....

---

---

---

---

---

---

---

---

## ISSN in SICI syntax



0886-9383(200112)15:10:1-K

Journal of chemometrics, 15 (2001), 10

The Italian ACNP catalogue  
processes SICIs for its holdings  
Co-managed by CNR Central Library and Bologna University

---

---

---

---

---

---

---

---

## identifying people ...

The new frontier of the identification

**ISNI,**

1422 4586 3573 0473

**ORCID**

0000-0001-7564-495X

The authors will be easier internationally identified

***Their 'grey literature products' will be a bit less grey ?***

---

---

---

---

---

---

---

---

# J-STAGE, NOW NEXT STAGE

**Full text database for reviewed academic papers published  
by Japanese Societies**

**More than 1,000 journals, 2 million records**

**80% full text at FREE**

**90% abstracts in English**

**Electronic Submission is acceptable for some journals**

**No registration is necessary**

■ <http://www.jstage.jst.go.jp>



# J-STAGE

JAPAN' S LARGEST PLATFORM  
FOR ACADEMIC E-JOURNAL



Gateway to a Wealth of Scientific and Technological Information  
Japan Science and Technology Agency(JST)

[www.jst.go.jp/EN](http://www.jst.go.jp/EN)

## Customized OAI-ORE and OAI-PMH Exports of Compound Objects for the Fedora Repository

Alessia Bardi, Sandro La Bruzzo, and Paolo Manghi

Istituto di Scienza e Tecnologie dell'Informazione,  
Consiglio Nazionale delle Ricerche, Italy

In order to address the needs of modern researchers to share and access research outcomes, Digital Library Systems (DLSs) evolved to manage new types of documents and collections which surpass the traditional publication-metadata document model to incorporate further entities involved in the research life-cycle. For example, *enhanced publications* [1] enrich the representation of traditional publications with other objects, such as metadata descriptions, research datasets, cited publications, etc. To cope with such representational requirements, modern Digital Library Management Systems (tools supporting DLS construction, ref. DELOS Digital Library Manifesto [2]) tend to offer *data models* capable of expressing DLS *document models* as “labeled graphs of information objects”. As a consequence, DLSs evolved also in the way they export their objects to third-party applications. Typically, objects are organized into “packages of interlinked information objects”, known as *compound objects*, and exposed through the standard APIs of the Open Archives Initiative protocols, which enable single or bulk access to compound objects respectively as OAI-ORE (RDF) aggregations [3] and OAI-PMH (XML) metadata records [4].

The Fedora Repository is a well-known DLMS platform [5], whose object data model is designed to represent arbitrary DLS document models by encoding them as graphs of *content model* objects, i.e., special Fedora objects declaring the “type of the objects” in a Fedora instance, i.e., the expected structure, relationships, and behaviors (Fedora methods) of the objects. This paper presents *OAIzer*, an extension of Fedora which is capable of exporting compound objects conforming to a given portion of the underlying DLS document model through the OAI-PMH or OAI-ORE protocols. The component implements a mechanism based on the concept of “*OAI view* of a Fedora document model”, that is a “sub-structure” of the document model which developers can provide to customize the shape of their compound objects. *OAIzer* interprets OAI views to automatically deploy web APIs capable of exporting compound objects compatible with the given structure and according to the preferred OAI protocol.

*OAIzer* is compared with other solutions for exporting compound objects in Fedora, namely *oreprovider*<sup>1</sup> and *Fedora2ORE*<sup>2</sup>. The former adopts an *object-oriented* approach, where ORE aggregations consist of sets of Fedora objects, the sets being identified by adding to the objects a pointer/reference to the relative ORE aggregation. The latter adopts instead an *object navigation-oriented* approach, where ORE aggregations consist of one Fedora object together with the objects reachable by navigating its relationships up to a given depth. The first solution is easy-to-use, but binds the aggregation logic to DLS applications and disregards document model relationships between objects. The second solution is independent from DLS applications and generates aggregations by following relationships between objects, but DLS developers can only define the boundaries of aggregations in terms of navigation depth rather than in terms of their preferred document model “sub-structure”. Finally, both solutions do not support OAI-PMH exports of compound objects.

---

<sup>1</sup> *OREprovider* Fedora module, <http://oreprovider.sourceforge.net/index.html>

<sup>2</sup> *Fedora2ORE*, <http://trac.eco4r.org/trac/eco4r/wiki/Fedora2ORE>

## References

1. Enhanced Publications, <http://www.surf.nl/en/themas/openonderzoek/verrijktepublicaties/Pages/default.aspx>, SURF Foundation
2. Candela, L., Castelli, D., Pagano, P., Thanos, C., Ioannidis, Y., Koutrika, G., Ross, S., Schek, H.J., Schuldt, H.: Setting the Foundations of Digital Libraries: The DELOS Manifesto. D-Lib Magazine 13 (2007)
3. Lagoze, C., Sompel, H. V. d., Johnston, P., Nelson, M., Sanderson, R., & Warner, S.. *Open Archives Initiative Object Reuse and Exchange*, <http://www.openarchives.org/ore/1.0/toc.html> (2008).
4. Lagoze, C., Van de Sompel, H.: *The open archives initiative: building a low-barrier interoperability framework*. In Proceedings of the first ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 54–62. ACM Press (2001).
5. C. Lagoze, S. Payette, E. Shin, C. Wilper, *Fedora: an architecture for complex objects and their relationships*, International Journal on Digital Libraries, pp. 124-138, Vol. 6, Issue 2, (2006).

## Bionotes

---

**Alessia Bardi** received her MSc in Information Technologies in the year 2009 at the University of Pisa, Italy.


She is a PhD student in Information Engineering at the Engineering Ph.D. School "Leonardo da Vinci" of the University of Pisa and works as graduate fellow at the Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" (ISTI), Consiglio Nazionale delle Ricerche (CNR) of Pisa, Italy. Alessia is currently involved in EC funded projects for the aggregation, curation and export of library, archival and museum digital objects and metadata records. Her research interests include Digital Library Management Systems, data interoperability, compound object management, and service-oriented infrastructures.

Email: [alessia.bardi@isti.cnr.it](mailto:alessia.bardi@isti.cnr.it)

**Sandro La Bruzzo** received his MSc in Information Technologies in the year 2010 at the University of Pisa, Italy. He is working as graduate fellow at the Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" (ISTI), Consiglio Nazionale delle Ricerche (CNR) of Pisa, Italy. His research interests include Digital Library Management Systems, compound object management, and Service-oriented Data Infrastructures.

**Paolo Manghi** received his PhD in the year 2002 from the Dipartimento di Informatica of the University of Pisa, Italy. He is presently working as a researcher at the Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo", Consiglio Nazionale delle Ricerche (CNR) of Pisa, Italy. His research interest include Data Models for Digital Library Management Systems, Types for Compound Objects, data curation in Digital Libraries, service-oriented ICT infrastructures with special focus on data ICT infrastructures.


Email: [Paolo.Manghi@ISTI.CNR.IT](mailto:Paolo.Manghi@ISTI.CNR.IT)



**CUSTOMIZED  
OAI-ORE AND OAI-PMH  
EXPORTS  
OF COMPOUND OBJECTS FOR  
FEDORA REPOSITORIES**

Alessia Bardi, Sandro La Bruzzo, Paolo Manghi  
name.surname@isti.cnr.it  
<http://nemis.isti.cnr.it/groups/infrascience>

Fourteenth International Conference on Grey Literature - Rome, 2012




---

---

---

---

---

---

---

---

## SCENARIO AND MOTIVATIONS

- Digital Library Systems (DLSs) need to export content
  - Export of compound objects: packages of information objects with an identity
  - OAI protocols (typical solution)
- DL management systems (DLMSs) need to support export protocols
- DLMS issues:
  - Absence of OAI-exports: some DLMS do not provide support for OAI protocols (Relational Databases)
  - Pre-defined OAI exports of compound objects: shape of objects as exported by OAI protocols is pre-defined by the DLMS (Fedora)

Fourteenth International Conference on Grey Literature - Rome, 2012

---

---

---

---

---

---

---

---

## OBJECT EXCHANGE PROTOCOLS

	OAI-PMH	OAI-ORE
Exchange subject	Metadata	Resource
Graph-based data model	No	Yes
Data Model	Set Item MetadataPrefix	Aggregation Aggregated Resource Proxy
Protocol	6 OAI Verbs on HTTP	HTTP
Data Serialization	XML	RDF (XML, ATOM)
Deleted records management	Yes	No

Fourteenth International Conference on Grey Literature - Rome, 2012

---

---

---

---

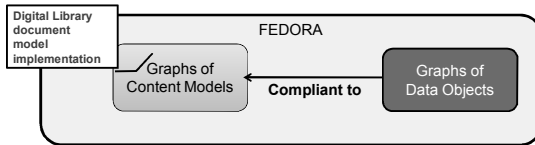
---

---

---

---

## FEDORA: FLEXIBLE EXTENSIBLE DIGITAL OBJECT REPOSITORY ARCHITECTURE



A Fedora instance manages graphs of Data Objects

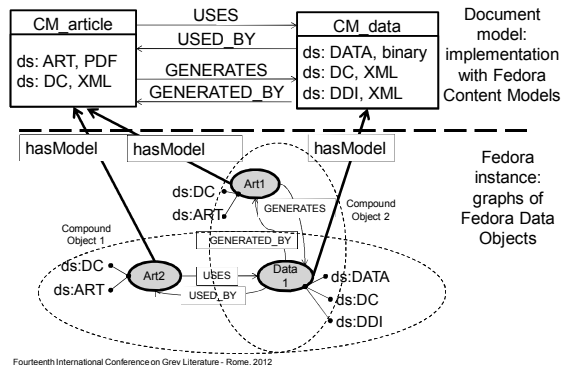
Data Objects conform to Content Models

Content Models define:

- Type and name of mandatory datastreams (files)
- Allowed relationships to other Data Objects

Fourteenth International Conference on Grey Literature - Rome, 2012

## FEDORA OBJECTS AND COMPOUND OBJECTS



Fourteenth International Conference on Grey Literature - Rome, 2012

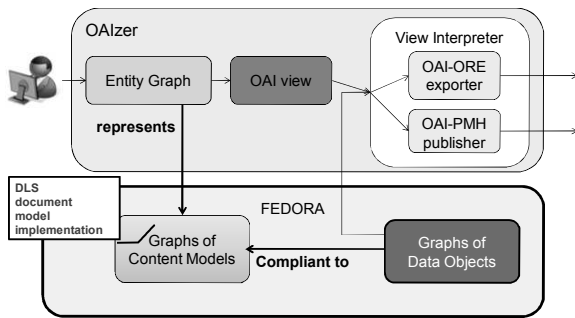
## EXPORTING COMPOUND OBJECTS IN FEDORA – SOLUTIONS AND LIMITS

- Current solutions for OAI-PMH exports on Fedora...
  - ...export OAI-items with the granularity of a Fedora Data Object
  - ...support the export of XML contained in datastreams
- Current solutions for OAI-ORE exports on Fedora...
  - ...exclude relationships and their semantics
  - ...allow limited customization of the structure of the compound objects (e.g., navigation depth)
  - ...pre-define the structure of the exported objects
- OAI-PMH works on any Fedora instance with Content Models:
  - ORE exports include relationships and their semantics
  - Structure of exported objects is fully customizable
  - OAI-PMH metadata formats can be generated on request (requires the configuration of an appropriate transformation rule)

Fourteenth International Conference on Grey Literature - Rome, 2012



## OAIZER: HIGH-LEVEL ARCHITECTURE

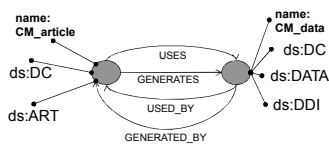


Fourteenth International Conference on Grey Literature - Rome, 2012

## PHASE 1: ENTITY GRAPH

Entity Graph: Graph representation of the document model

- Created exploiting information about Content Models
- Graph nodes represent Content Models
- Properties of a node represent the datastreams defined by the Content Model
- Edges between nodes represent the relationships defined by the Content Models

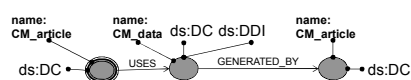


Fourteenth International Conference on Grey Literature - Rome, 2012

## PHASE 2: DEFINITION OF THE OAI VIEW

OAI View: Tree representation of the structure of the compound objects to export

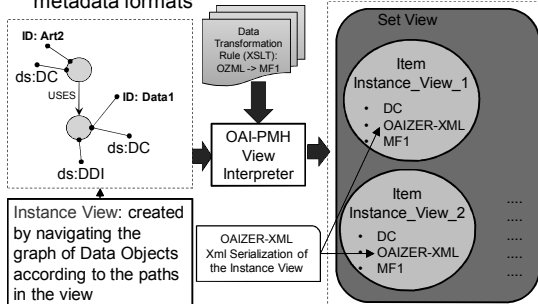
- Sub graph of the entity graph with entities and relationships to include in the compound object
  - Ex. The DATA datastream of CM\_data is excluded
- The root node is the entry point of the OAI view



Fourteenth International Conference on Grey Literature - Rome, 2012

### PHASE 3.1: INTERPRETATION OF THE OAI VIEW FOR OAI-PMH

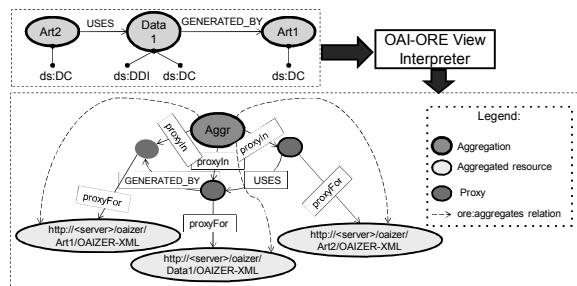
- Compound object represented as OAI-Item with several metadata formats



Fourteenth International Conference on Grey Literature - Rome, 2012

### PHASE 3.2: INTERPRETATION OF THE OAI VIEW FOR OAI-ORE

- Compound object exported as OAI-ORE Aggregations.
- Aggregated resources in OZML format



Fourteenth International Conference on Grey Literature - Rome, 2012

### CONCLUSIONS

- OAIzer's wizards guide DLSs developers through the realization of customized, domain-dependent exports of compound objects via OAI-PMH and OAI-ORE
- Future work:
  - Better graphical support for the definition of views in case of entity graphs with cycles
  - Implementations for different back-ends
    - RDBMS: ongoing work

Fourteenth International Conference on Grey Literature - Rome, 2012

# Grey literature in the digital culture and practices of the new global scholar: the case of molecular biology

Chérifa Boukacem-Zeghmouri

Département d'Informatique de l'Université Claude Bernard Lyon 1, France

## Problem

Molecular biology positions itself as a cutting-edge field within the new globalized modalities of scientific communication, which is characterized by data sharing, collaborative work, all supported by a distributed IT infrastructure. The recent announcement of the new generation of high-speed sequencers of the human genome, for the end of 2012, reinforces the field's cutting-edge position. This method – genome sequencers – will be used to conduct genetic tests for cancer research.

The digital culture and practices of molecular biologists have been studied, given their innovative nature (Gallezot, 2002) (Morris, 2005). These digital practices are related to Big Science, indicating a paradigm change in the practice of science itself. However, while studies have been conducted on the structuring practices of molecular biology, they haven't focused on grey literature.

We know that researchers in molecular biology rally research data, genome information, publications and many other sources of knowledge. That being said, the importance and place of grey literature in the process of knowledge production remain unknown. This is precisely what we will investigate in the context of this work.

## Methodology

Our methodology rests on qualitative research based on a dozen semi-directive interviews with researchers in molecular biology, paired with an ethnographic observation. The goal is to identify the share of grey literature in the researchers' informational habits and behaviour.

We are containing our study to the CRCL French research laboratory<sup>1</sup>, in Lyon. The lab's size (350 people), the recognized quality of its research and the plurality of studied topics allow us to investigate our own questions, while always taking into account the specific context – the CRCL laboratory – in which our research takes place.

## Expected results

Analyzing the interviews will not only help us to identify the amount of grey literature present in the information practices of researchers in molecular biology. We will also be able to know what kind of grey literature these researchers produce and distribute. In the end, we will be able to estimate grey literature's specific contribution to the research life cycle and to the researchers' innovation process.

## Estimated costs

25 days = approximately. 5,000.00 EUR (Salary). Transportation costs : 53 EUR.

## Bibliography

- W. Y. Arms, et al. (2009). 'Escience in Practice: Lessons from the Cornell Web Lab'. *D-Lib Magazine* 15(5/6).
- C. L. Borgman (2007). *Scholarship in the digital age : information, infrastructure and the Internet*. MIT Press.
- J. Fry, et al. (2009). 'Open science in e-science: contingency or policy?'. *Journal of Documentation* 65(1):6–32.
- T. Gabridge (2009). 'The Last Mile: Liaison Roles in Curating Science and Engineering Research Data'. *Research Library Issues: A Bimonthly Report from ARL, CNI and SPARC* (265):15–21.
- S. A. Gore (2011). 'e-Science and Data Management Resources on the Web'. *Medical Reference Services Quarterly* 30(2):167–177.
- G. Gallezot (2002a). 'Exploration informationnelle et construction des connaissances en génomique'. *Les Cahiers du numérique* pp. 121–136.
- G. Gallezot (2002b). 'La recherche in silico'. *Les chercheurs et la documentation électronique: nouveaux services, nouveaux usages*. Paris: Édition du Cercle de la Librairie .
- T. Hey, et al. (eds.) (2009). *The fourth paradigm. Data-intensive scientific discovery*. Microsoft Corporation.
- S. MacKenzie (2011). 'Communicating with data: new roles for scientists, publishers and librarians'. *Learned Publishing* 24(3):203–205.
- R. W. Morris, et al. (2005). 'Digital Biology: an emerging and promising discipline'. *Trends in Biotechnology* 23(3):113–117.
- A. Renear & C. Palmer (2009). 'Strategic reading, Ontologies, and the future of scientific publishing'. *Science* 325:828–832.
- S. B. Weilandt (2007). 'E-Science à "Advancing new ways of scientific communication."'. *Information Services & Use* 27(4).

<sup>1</sup> - <http://www.centreonberard.fr/74-Centre-de-recherche-en-cancerologie.clb.aspx>



## Bionote

---

**Chérifa Boukacem-Zeghmouri** is Lecturer in Information and Communication Science at Université de Lyon (Lyon 1 University – Sciences). She is a research member of the ELICO Research Laboratory. Her PhD has been dedicating to an economic analysis of the transition of the academic library services in an electronic context (defended in 2004). From 2006 to 2010, Chérifa Boukacem-Zeghmouri chaired a national project on the usage of electronic resources in the French academic context, using a socio-economic research approach. She is now coordinating a national study on the ROI of electronic resources in the French academic network. She has been member of the Research Oversight Group of the PEER Project. Since October 2011, she is Head of the Regional Unity of Training to Scientific Information dedicated to scholars, PhD students, librarians and information professionals (URFIST Lyon). Email: [cherifa.boukacem@univ-lyon1.fr](mailto:cherifa.boukacem@univ-lyon1.fr)

## Grey in the Innovation Process

**Keith G Jeffery**, STFC-Rutherford Appleton Laboratory, United Kingdom

**Anne Asserson**, University of Bergen, Norway

The research lifecycle has multiple objectives materialised as outputs, outcomes and impacts. Typical outputs are research publications (including grey literature), patents and products such as research datasets and software, many kinds of art or prototype engineering artifacts. Outcomes include patent licence income, value of a company set up to exploit the output or trained research staff. Impacts include employment creation, a commercial product that saves lives or labour or development of a new field of knowledge and research such as genomics since the 1950s.

Commonly research in progress may be documented as grey literature – such as technical reports, laboratory notebooks or instructions for operating new equipment. There is a decision point when grey literature is produced.

One can innovate academically. The output is peer reviewed publications; the outcomes include developing trained researchers; the impact leading to a new field of research. This route provides academic recognition.

Alternatively one can innovate along the wealth-creation route. The output could be a patent; the outcome license income or a new company; the impact employment, dividends to shareholders or a new 'wonder product'. This route provides wealth and possibly improvement in the quality of life.

If research is published this usually precludes following the wealth-creation route since the novel idea is now in the public domain and not protected by patent(s).

Increasingly research funding and research performing organisations wish to demonstrate that the research they fund or do leads to impacts of relevance to society. Tracing of impacts back to the original research is not easy, partly because the eventual impact may not be known for many years. The key is an accurate recording of the research lifecycle including important dates so that the innovation cycle from idea to impact and back to further ideas can be demonstrated.

Recent work – especially in UK in the JISC-funded MICE project – has produced a taxonomy of outputs, outcomes and impacts. In parallel an extension to CERIF (Common European Research Information Format – an EU recommendation to member states) has been developed and approved by euroCRIS. This extension re-uses typical CERIF entities of persons, projects, organisations, publications, patents, products but relates them (with temporal validity and appropriate role) to the production or utilisation of outputs, outcomes and impacts. Naturally grey literature is a key component within this model.

## Bionotes

**Keith Jeffery** is currently Director International Relations at STFC (Science and Technology Facilities Council) based at Rutherford Appleton Laboratory. Keith previously had strategic and operational responsibility for ICT with 360,000 users, 1100 servers and 140 staff. Keith holds 3 honorary visiting professorships, is a Fellow of the Geological Society of London and the British Computer Society, is a Chartered Engineer and Chartered IT Professional and an Honorary Fellow of the Irish Computer Society. Keith is currently President of ERCIM and President of euroCRIS, and serves on international expert groups, conference boards and assessment panels. He has advised government on security and green computing. He chaired the EC Expert Groups on GRIDs and on CLOUD Computing.

Email: keith.jeffery@stfc.ac.uk

**Anne Asserson** holds a master from the University of Bergen, UiB. She has been working with Research Documentation, and has participated in substantial parts of CRIS developmental work, locally and nationally since 1992. Anne Asserson has been part of the establishing and implementing of several CRIS both at the UiB and nationally. For several years she was the chairwoman of the Steering Group of the national CRIS system and project secretary of a National system for academic administration. Anne Asserson is presently representing UiB in the national project group of CRISin. She has also participated in The CORDIS funded European-wide project on " Best Practice" 1996 and was a member of the working group set up 1997 that produced the report CERIF2000 Guidelines (1999) [www.cordis.lu/cerif](http://www.cordis.lu/cerif), coordinated by the DGXIII-D4. euroCRIS is now the custodian of the CERIF model [www.eurocris.org](http://www.eurocris.org). Anne Asserson is a member of the euroCRIS board with the responsibility Member Strategy and External Relations. Email: anne.asserson@fa.uib.no

# Grey in the Innovation Process

Keith G Jeffery  
Director, IT &  
International Strategy, STFC

Anne G S Asserson  
University Library  
University of Bergen

keith.jeffery@stfc.ac.uk

anne.asserson@fa.uib.no



© Keith G Jeffery, Anne Asserson

Grey in the Innovation Process

29-30 Nov 2012 Rome

1

---

---

---

---

---

---

---

---

## Authors

Keith G Jeffery  
STFC-RAL



Anne Asserson  
UiB

© Keith G Jeffery, Anne Asserson

Grey in the Innovation Process

29-30 Nov 2012 Rome

2

---

---

---

---

---

---

---

---

## Structure

- Innovation
- Background
- The Hypothesis
- Proposed Architecture
- Conclusion



"I'll be happy to give you innovative thinking. What are the guidelines?"



© Keith G Jeffery, Anne Asserson

Grey in the Innovation Process

29-30 Nov 2012 Rome

3

---

---

---

---

---

---

---

---

## Innovation

**Innovation** is the development of new customers value through solutions that meet new needs, inarticulate needs, or old customer and market needs in new ways.  
(Wikipedia)

→  
IDENTIFY NEW NEEDS  
DEVELOP NEW WAYS



© Keith G Jeffery, Anne Asserson

Grey in the Innovation Process

29-30 Nov 2012 Rome

4

## Background

- Need to justify research funding:
  - Wealth creation
  - Improvement in quality of life
- Through Innovation
- Industry R&D
  - Supported by government
- Government R&D
  - Supported by industry
- Research spend
  - Gross R&D spend
- Research intensity
  - R&D spend as % of GDP

© Keith G Jeffery, Anne Asserson

Grey in the Innovation Process

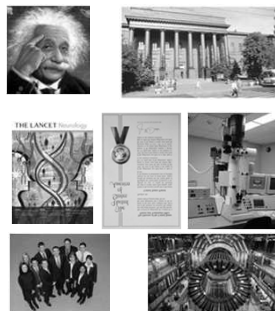
29-30 Nov 2012 Rome

5

## The Hypothesis: Understanding innovation requires:

- Measures
  - Outputs
  - Outcomes
  - Impacts
- Contextual relationships
  - Funding
  - Institution
  - Project
  - Person
  - Facility / Equipment
- Temporal relationships
  - Provenance
- Geospatial relationships

→ Causality



© Keith G Jeffery, Anne Asserson

Grey in the Innovation Process

29-30 Nov 2012 Rome

6

## The Hypothesis: GREY is part of this

- Documented original idea
  - Lab notebook
  - Internal report
- Document process of development
  - Lab notebook
  - Internal report
  - Workshop presentation
- Management control of IP



© Keith G Jeffery, Anne Asserson

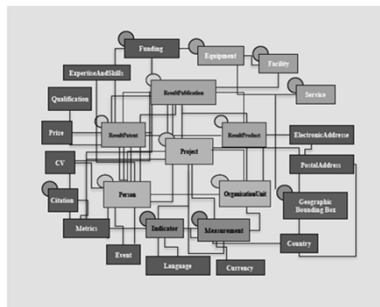
Grey in the Innovation Process

29-30 Nov 2012 Rome

7

## Proposed Architecture: CERIF

- Describe
  - Outputs
  - Outcomes
  - Impacts
- Record
  - Contextual relationships
  - Temporal relationships
  - Geospatial Relationships
- Deduce or Induce
  - Causality



© Keith G Jeffery, Anne Asserson

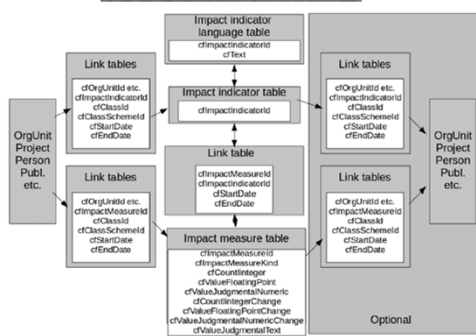
Grey in the Innovation Process

29-30 Nov 2012 Rome

8

## MICE

Measuring Impact under CERIF



© Keith G Jeffery, Anne Asserson

Grey in the Innovation Process

29-30 Nov 2012 Rome

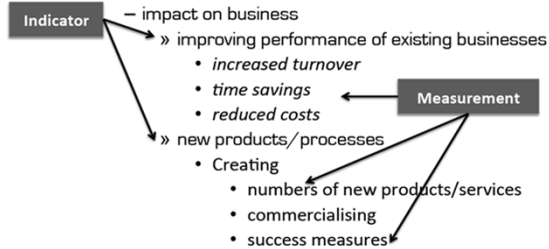
9



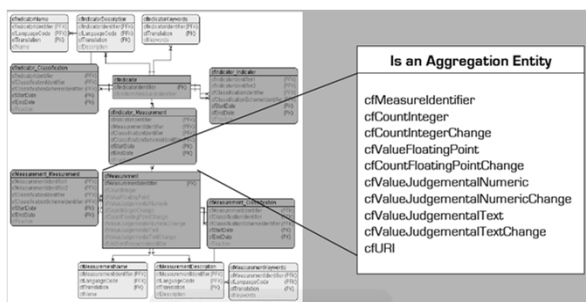
## Example of Commercial Innovation Indicators and Measures

– economic and commercial

- economic



## CERIF Data Structure



## Conclusion

- CERIF provides data structure for innovation
  - Including within the GREY process
- It is being used in significant systems
  - Tracking outputs, outcomes, impact
  - Related to contextual, temporal, geospatial metadata
- euroCRIS has an Indicators Task Group
  - New scientometrics (includes bibliometrics)
  - New methods for detecting impact (backward chaining)

- Acknowledgements
  - euroCRIS
  - Especially MICE project



## **Characteristics and use of grey literature in scientific journals articles of Algerian University of Science and Technology teachers and researchers at STM fields**

**Lydia Chalabi,**

Research Center of Scientific and Technical Information, CERIST, Algeria

Scientific research includes research activities, for creating new knowledge in the form of research output. This production search are governed either by a formal mode of communication and trade, such as, books, scientific journal articles, or an informal mode of communication as theses, conference proceedings and research reports that is called grey literature. However, with new technologies and the advent of open access to information, production and dissemination of scholarly publications, knows a revolution: diversity, speed and wide availability. In addition, nowadays, a strong need to evaluate the effectiveness of basic research appeared. Scientific research must meet the current modes of communication and understanding new and emerging needs, such examples, a global and large communication and competition between researchers, institutions, countries.

In Algeria, despite the inadequacies of financial resources, documentaries, and the presence of some constraints, like economic, impeding the realization of research works, academic institutions and research have not escaped the Algerian context of globalization.

Indeed, the latest results of the evaluation of research institutions and ranking Algerian emerging disciplines (for Thomson Reuters in January 2012) reveal the presence, although limited, of the scholarly output of researchers from Algeria in the prestigious international databases, which is encouraged to determine the incentives of that and support it.

In this respect, our paper focuses on the presence and characteristics of the grey literature in the scientific literature, a pillar of scholarly communication, of Algerians teachers and researchers at Algerian University of Science and Technology Houari Boumediene (USTHB) first Algerian university in terms of publications (Thomson Reuters in January 2012) and in the most emerging STM( Science, Technology and medicine covering applied science) field, governed by the traditional system of scientific publishing, ie commercial, based on the following issues:

- Algerian researchers from the University USTHB, use they a grey literature in their scientific articles?
- How are they informed by the presence of grey literature?
- What are the different sources of dissemination of grey literature they use? Libraries? web resources? Commercial or free?
- Use they the open repositories of grey literature?
- Can we argue that open access has affected research and foster its development through more scientific publication?

### **Interest and objective of the study**

This study is the first of its meaning, it can be seen as a contribution to stimulate interest in an issue that we consider important to the Algerian researchers, both concerns in the assessments of research institutions to decide how to orient and improve their acquisitions, information policy, and information dissemination. It also contributes to study the impact of open access on publication of Algerian researchers.

### **Hypothesis**

To achieve these objectives, we made two assumptions that will form the basis for the collection of information:

- Grey literature is used in the research of Algerian researchers, particularly in scientific articles.
- The Algerian researchers use, much more, the open access grey literature retrieved on the web.



### Methodology

Our methodology is based on a qualitative study (semi-structured interviews) with Algerian researchers in the field of STM (Physic, biology and chemistry) USTHB University (ranked first Algerian university in terms of greater number of scientific publication by Thomson Reuters).

### Bionote

---

**Lydia Chalabi** is a researcher at Research and Development in Information Sciences Division in Algerian Research Center on Scientific and Technical Information (CERIST). She studied Library Sciences at university of algeirs2 where she obtained her first graduate diploma by studying the on-line Algerians scientific journals. And a Master's Degree related to the provision of on-line bibliographic services: a case study of Algerian books. Since 2010, she is preparing a PhD in information sciences about the impact of open access on the Algerians researcher's science. Since 2009, she teaches the documentation networks course in the CERIST e-learning platform. For 2006 to now, she held a various research activities and projects in the field of information sciences and currently, she is a research project leader concerning the impact of open access on an Algerian scientific literature. Her research special fields of interests are on Open Access to information sciences, open archives and open repositories, e-publishing, impact factor, open access citation impact and electronic resources. Email: [net\\_lydia@yahoo.fr](mailto:net_lydia@yahoo.fr)

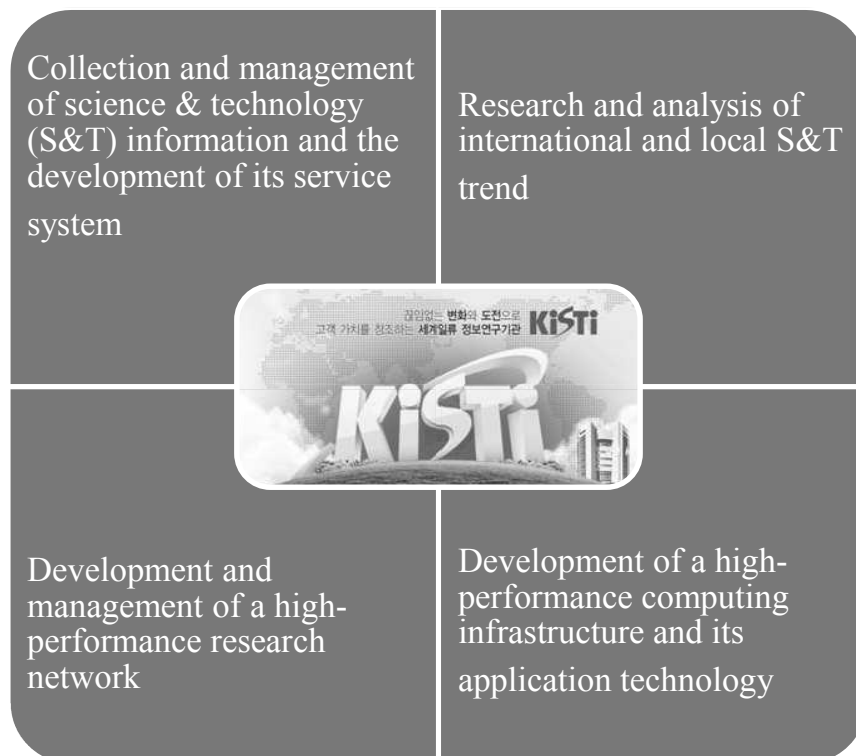
# Korea Institute of Science and Technology Information (KISTI)

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

## \* Vision

World-class information research institute creating values for customers

## \* Main functions



## \* 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](mailto:hcpark@kisti.re.kr)

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

## The Research Life Cycle and Innovation through Grey Literature in Nanotechnology in Korea

**Seon-Hee Lee,**

*Korea Institute of Science and Technology Information, KISTI, Korea*

This paper studied the research life cycle of Nanotechnology (NT) and its innovation through grey literature in Korea. The goals of this study are to analyze research life cycle in Nanotechnology and to trace of innovation through grey literature (technical reports). I assumed changes of numbers of publications of grey literature and white literature in NT show the process of innovation because technical reports, journal articles contained research results. Numbers of publications of Korean technical reports and journal articles on NDSL and Web of Science are compared year by year.

For analysis of the research life cycle in NT in Korea, we conducted an in-depth interview and close observation were conducted with 24 researchers from March to April in 2011 in the field of nanotechnology. The research life cycle consisted of idea building, funding, experiment and analysis, result creation, and evaluation. The researchers use the grey literatures in every stage of their research cycles but also produced them. As result of creation, technical reports, journal articles, patents are conducted.

The scholarly communication system in Korea is mentioned by National Science & Technology Information System (NTIS) and National Discovery for Science Leaders (NDSL). The research results by scholars and researchers have been collecting as the forms of the technical reports of national research and development projects through NTIS and the journal articles through Article Contribution Management System (ACOMS). The researchers are supposed to upload their technical reports and Journal articles on the systems via the internet by themselves. Then the grey literatures and white literature uploaded are being serviced to public at a portal system so called NDSL.

Technical reports funded by Korean government on NT have been published since 1980'. The peak of Technical reports publication was researched in 2005 and then reduced. The early stage of research development of NT has appeared in the grey literature. The numbers of domestic journal articles increase almost same as technical reports. Numbers of Web of Science (SCIE) journal articles written by Korean have been increasing dramatically since 2005. 2005 is the turning point for dominant literature in Nanotechnology from grey literature to white literature. Nanotechnology was the 10<sup>th</sup> most published subjects in WoS journal articles written by Korean in 2011. Nanotechnology became one of the 6 technologies supported by Korean government and 2 most funded technology by Korean government. Technical reports as grey literature provide creative idea and research results in early stage of development of NT in Korea. These results indicated that innovation in NT has taking place through grey literature.

### Bionote

---

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



## The Research Life Cycle and Innovation through Grey Literature in Nanotechnology in Korea

Nov. 29, 2012

Seon-Hee Lee

Korea Institute of Science and Technology Information

---

---

---

---

---

---

---

## Contents

- Introduction
- Research Life Cycle in Nanotechnology in Korea
- Scholarly Communication System in Korea :  
NDSL, NTIS
- The Comparison of Technical Reports and  
Korean Journal Articles on NDSL and WoS (SCIE)  
written by Korean
- Trace of Innovation in Technical Reports and  
Articles in WoS in Nanotechnology

---

---

---

---

---

---

---

## Introduction

### Goal of study

**To analyze research life cycle in Nanotechnology(NT)**

**To trace innovation through grey literature in Korea**

- **Assumption** : Changes of numbers of publications of grey literature and white literature in NT show the process of innovation because technical reports and journal articles contained research results.
- **Data** : Numbers of publications of Korean technical reports and journal articles in NT on NDSL and Web of Science
- **Comparison** : Numbers of publications as research results in technical reports and journal articles year by year

---

---

---

---

---

---

---

## Needs and Goal of Analysis of Research Life Cycle in Korea

### Needs of R&D research life cycle

- As national information center for Science and Technology, KISTI should know domestic researchers information environment and provide stable system.
- Conducted Researchers' R&D life cycle, to adjust and to prepare for changeable information environment for users is important.

### Goal of R&D research life cycle in nanotechnology

- What is R&D research life cycle in nanotechnology in Korea

---

---

---

---

---

---

---

---

## Research Life Cycle in NT

### Conduct interview

- Method : In-depth interview and close observation
- Interviewee : 24 researchers in Nanotechnology
- Interview Period : March – April in 2011 (2 months)

### Analysis of research life cycle

- Divided into **5 Stages** : idea building, funding, experiment and analysis, result creation, and evaluation
- Needs are different in every stage of research life cycle
- **Need literature reviews (technical reports, patents, trends , journals) using NDSL NTIS and etc. in every stage of research life cycle**
- **Research results in grey and white literature, industrialization**

---

---

---

---

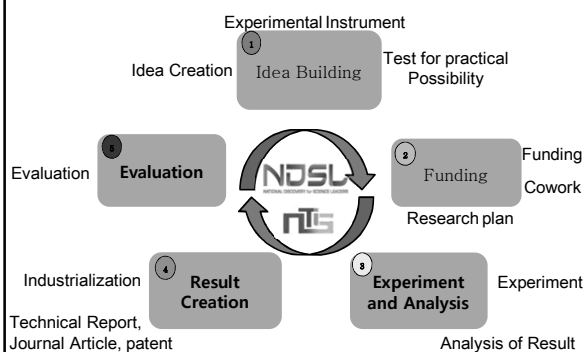
---

---

---

---

## Research Life Cycle in NT




---

---

---

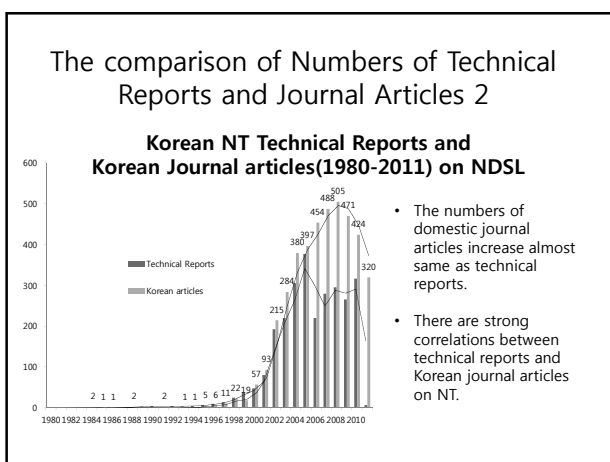
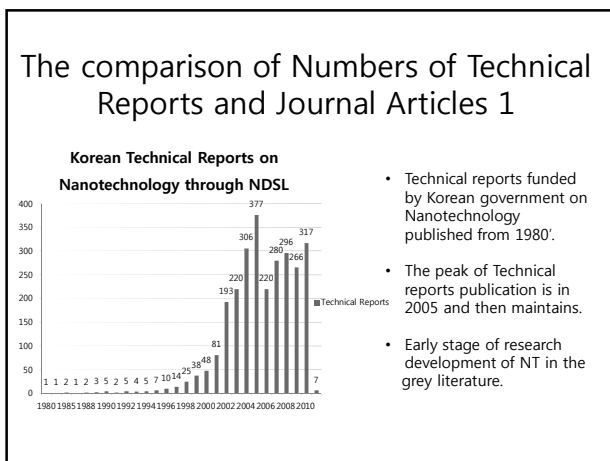
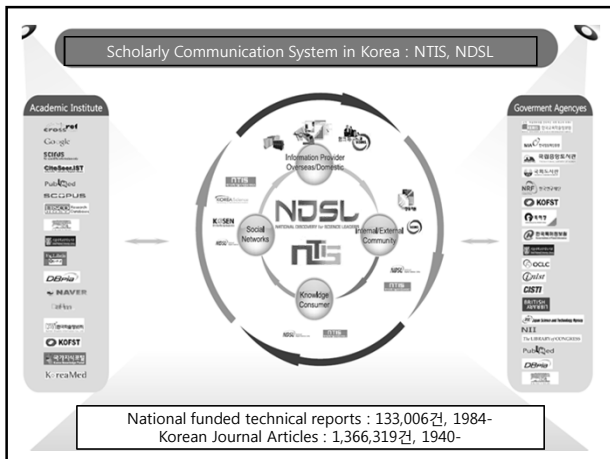
---

---

---

---

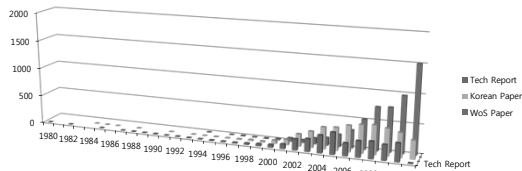
---





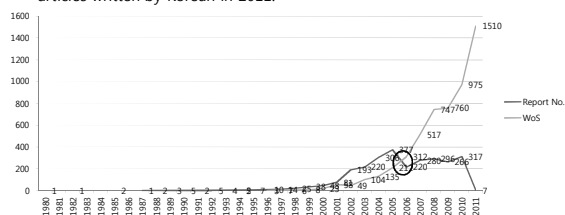
### The comparison of Numbers of Technical Reports and Journal Articles 3 : WoS (SCIE)

- Numbers of technical reports and domestic journal articles are reducing since 2005.
- WoS journal articles written by Korean have been increasing dramatically.



### The comparison of Technical Reports and Journal Articles 4

- 2005 is the turning point for dominant literature in NT.
- => from Technical reports to Journal articles on WoS
- WoS journal articles written by Korean has increased dramatically.
- Nanotechnology is the 10<sup>th</sup> most published subject for WoS journal articles written by Korean in 2011.



### Trace of Innovation through Technical Reports in Nanotechnology

- Technical reports as grey literature provide creative idea and research results in early stage of development of NT in Korea.
- The Innovation in NT
  - Numbers of publication are increased in Web of Science(SCIE),
  - the 10<sup>th</sup> most published SCIE papers written by Korean in 2011.
  - NT emerged in 1980' and became one of the 6 technologies supported by Korean government.
  - NT was the 2 most funded technology by Korean government
- Innovation in NT has taking place through grey literature.**

**Thank you!!**

# *Library, Information Science & Technology Abstracts<sup>TM</sup> with Full Text*

*Available via EBSCOhost<sup>®</sup>*

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

*Library, Information Science & Technology Abstracts<sup>TM</sup> with Full Text* is an indispensable tool for librarians looking to stay current in this rapidly evolving field.

**Comprehensive content includes:**

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

**Subject coverage includes:**

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

Contact EBSCO Publishing to learn more about *Library, Information Science & Technology Abstracts<sup>TM</sup> with Full Text*, or to request a free trial.

Phone: 800.653.2726

Email: [request@ebscohost.com](mailto:request@ebscohost.com)

[www.ebscohost.com](http://www.ebscohost.com)

## What goes up must come down: Publications from developing countries in the Aquatic Commons

**Maria Kalentsits and Armand Gribling**

Fisheries & Aquaculture Branch Library, FAO  
Food and Agriculture Organization of the United Nations, Italy

FAO's Fisheries and Aquaculture Branch Library (FBL) is involved in a project which includes the selection, digitizing, web-optimization, creation of metadata and uploading into the Aquatic Commons (AC) of grey literature published by, amongst others, issuing agencies in several African countries and a regional project in Asia, the STREAM Initiative. Furthermore links to these full text online versions have been added to the Aquatic Sciences and Fisheries Abstracts (ASFA) bibliographic database.

The AC is a thematic digital repository covering the marine, estuarine, brackish and fresh water environments. The repository is set up by the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC) and hosted by the International Oceanographic Data and Information Exchange (IODE) project office in Belgium.

Together with EPrints software, the repository uses the related Interoperable Repository Statistics (IRS) software for usage statistics. In our paper we will use the IRS software to analyze the access to the grey literature from selected issuing agencies, whereby we will focus specifically on the Technical Documents, published by the Lake Victoria Fisheries Research Project and those published by the Support to Regional Aquatic Resources Management (STREAM), which was based in Bangkok, Thailand.

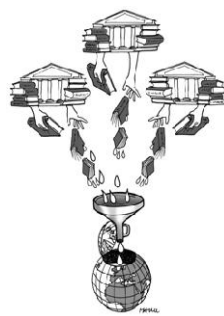
We will discuss examples of digital preservation and repatriation to countries of origin of documents from developing countries, two of the main objectives for FAO's participation in the repository.

As the repository was started in 2007 and the uploading of the papers was done only more recently, it is too early for citation analysis. By using the IRS software the paper seeks to address a further evaluation of FAO's contribution to the content development of the repository, confirming the increasing visibility of and access to some baseline information in the field of fisheries and aquaculture published by institutions and/or projects in developing countries. Finally the paper will describe some of the utilities and limitations of the software.

### Bionote

---

**Armand Gribling** is working in the Fisheries & Aquaculture Branch Library in FAO and is involved in the contents development of the Aquatic Commons repository, and member of the AC Board. He is co-author of a paper presented at the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC) Conference in 2010: Visibility and access through the Aquatic Commons (<https://darchive.mblwhoilibrary.org/handle/1912/4581>). Email: [armand.gribling@fao.org](mailto:armand.gribling@fao.org)



**WHAT GOES UP MUST COME DOWN:**

**Publications from developing countries in the Aquatic Commons**

Maria Kalentsits & Armand Gribling  
Fisheries & Aquaculture Branch Library  
Food and Agriculture Organization  
of the United Nations (FAO)

---

---

---

---

---

---


---

---

**WHAT GOES UP MUST COME DOWN: Publications from developing countries in the Aquatic Commons**

**Presentation Outline**

- Resource Sharing
- Aquatic Commons Digital Repository (AC)
- ASFA Trust Fund Project on Grey Literature for AC
- Statistics and Usage Analysis: Examples of Repatriation
  - Lake Victoria Fisheries Research Project (LVFRP)
  - Support to Regional Aquatic Resources Management Project (STREAM)
- Conclusions




---

---

---

---

---

---

---

---

**WHAT GOES UP MUST COME DOWN: Publications from developing countries in the Aquatic Commons**

**Resource Sharing**

**International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC)**

- Z39.50 Distributed Library
- Union List of Serials
- Duplicate Exchange

**International Oceanographic Data & Information Exchange Programme (IODE) of the Intergovernmental Oceanographic Commission IOC/UNESCO**

- OceanDocs
- Open Science Directory

**AQUATIC COMMONS**

**Fisheries and Aquaculture Branch Library (FBL) of the Food and Agriculture Organization of the United Nations (FAO)**

- ILL services
- Retrospective Digitization
- Corporate Document Repository

**Aquatic Sciences and Fisheries Abstracts (ASFA)**

- 1,651,725 bibliographic records (Oct. 2012)
- 66 co-sponsoring, national & international partners
- ASFA Thesaurus

---

---

---

---

---

---


---

---

WHAT GOES UP MUST COME DOWN: Publications from developing countries in the Aquatic Commons

### Aquatic Commons

<http://aquaticcommons.org/>



- Covers the natural marine, estuarine/brackish, and fresh water environments
- Established in 2007, directed by IAMSILIC, hosted by the UNESCO/IOC Project Office for IODE
- Powered by EPrints, running version 3.2.9
- 93 official issuing agencies
- ~ 7300 uploaded documents
- >500,000 downloads

---

---

---

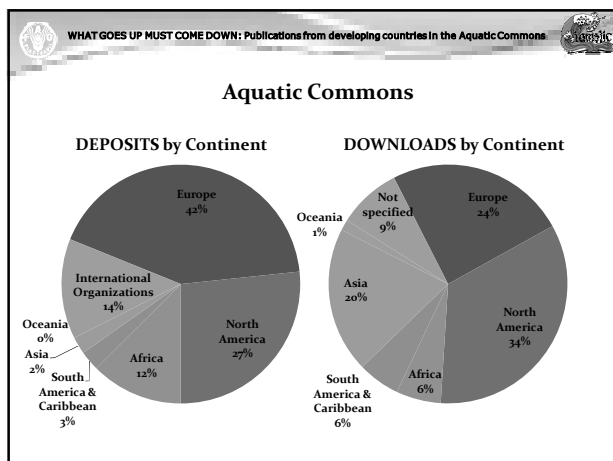
---

---

---

---

---




---

---

---

---

---

---

---

---

WHAT GOES UP MUST COME DOWN: Publications from developing countries in the Aquatic Commons

### Aquatic Sciences and Fisheries Abstracts - Trust Fund

Small projects aiming at increasing visibility of grey literature from partner institutions :

- Through filling gaps in the ASFA bibliographic database
- By digitization of grey literature and adding full text links to ASFA records

**ASFA Trust Fund Project "Published and Grey Literature of African Institutions for Aquatic Commons"**

- Proposed by IAMSILIC and completed by FBL
- More than 700 documents digitized and uploaded
- Links to full text documents in ASFA records
- 7 issuing agencies




---

---

---

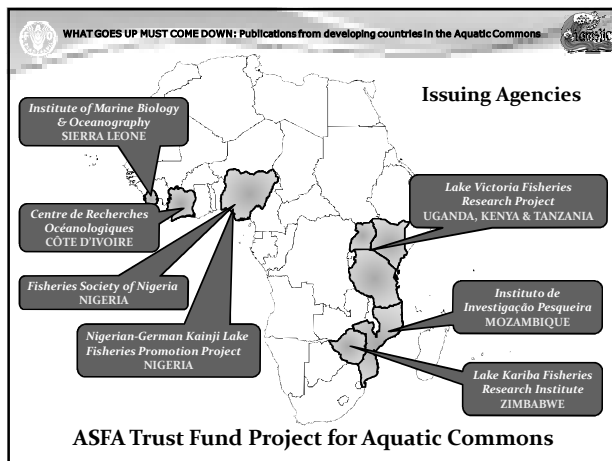
---

---

---

---

---




---

---

---

---

---

---

---

---

WHAT GOES UP MUST COME DOWN: Publications from developing countries in the Aquatic Commons

**LVFRP I: the project**

- The Lake Victoria Fisheries Research Project: framework for the management of the Lake's fisheries and a basis for a Fisheries Management Plan, which has been adopted by the Council of Ministers of the Lake Victoria Fisheries Organization (LVFO)
- LVFRP Technical Documents (75 AC records)

<http://www.lvfo.org/>

---

---

---

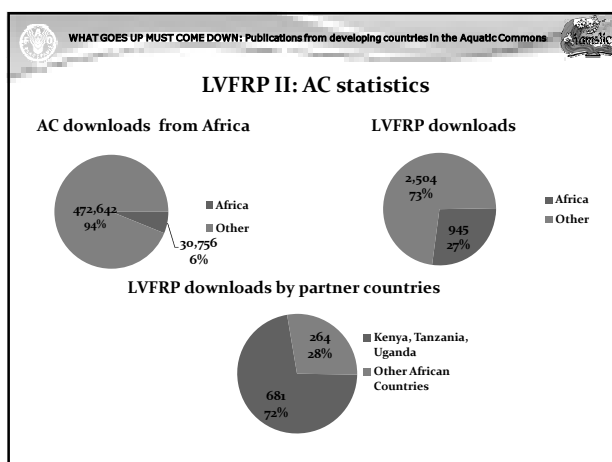
---

---

---

---

---




---

---

---

---

---

---


---

---

WHAT GOES UP MUST COME DOWN: Publications from developing countries in the Aquatic Commons

### STREAM I: the project

- Support To Regional Aquatic Resources Management: promoting participation, communication, and policies that support the livelihoods of poor aquatic resource users in the Asia-Pacific
- Example of *preservation*; project website up from 2002-2008
- At present the selected 126 documents almost exclusively available from Aquatic Commons




---

---

---

---

---

---

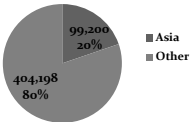
---

---

WHAT GOES UP MUST COME DOWN: Publications from developing countries in the Aquatic Commons

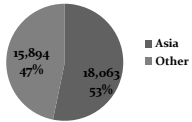
### STREAM II: AC statistics

AC downloads from Asia



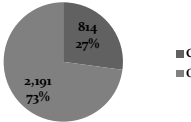
Region	Count	Percentage
Asia	404,198	80%
Other	99,200	20%

STREAM downloads from Asia



Region	Count	Percentage
Asia	15,894	47%
Other	18,063	53%

Downloads of STREAM documents on Cambodia



Region	Count	Percentage
Cambodia	814	27%
Other	2,191	73%

---

---

---

---

---

---

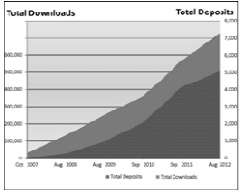
---

---

WHAT GOES UP MUST COME DOWN: Publications from developing countries in the Aquatic Commons

### Conclusions

- Digitization of grey literature published by fisheries and aquaculture institutions in developing countries:
  - Increases overall **access**
  - Contributes to **preservation**
  - Enables "**repatriation**" to countries of origin
- IRS software can be applied for AC usage analysis
- FAO's Fisheries and Aquaculture Branch Library will continue to promote the participation of institutions in developing countries in the Aquatic Commons




---

---

---

---

---

---

---

---

## **Data sharing in environmental sciences: A survey of CNR researchers**

**Daniela Luzi**

Istituto di Ricerche sulla Popolazione e le Politiche Sociali, IRPPS-CNR, Italy

**Roberta Ruggieri**

Senato della Repubblica, Italy

**Stefania Biagioni**

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

**Elisabetta Schiano**

Institute of Marine Sciences, ISMAR-CNR, Italy

Today the free availability of research data is considered an important driver of innovation and of new scientific insights. Due to the increasing amount of data collected as well as to the variety of purposes, process of acquisition and formats this is not an easy task. It implies the development of policies that incentive data curation and preservation, the recognition of the value of research data as “first-class publication”, the enforcement of clear rules for open access, copyright and ownership. It is also necessary that the scientific community agree on the development and use of common interoperability standards related to data models, format and exchange protocols. Last but not least it requires that suitable infrastructures be developed at national and international level considering discipline specificity.

Projects aiming to develop infrastructures for research data as well as metadata standards that enable data sharing are carried out in many disciplinary and cross-disciplinary fields. Moreover, many surveys are exploring scientists’ practices and perceptions toward data acquisition, curation and preservation focusing in particular on perceived barriers and enablers of data sharing.

The paper presents a survey based on an online questionnaire submitted to CNR researchers in the area of Environmental Sciences that represent a data intensive, collaborative and multi-disciplinary field. The survey has the aim to explore researchers’ attitude in data management, use and sharing, considering practices employed in all phases of data lifecycle as well as needs and requirements that are felt to be crucial for an effective implementation of data preservation.

The questionnaire consists of two sections. The first one provides the demographic characteristics of the sample and addresses questions related to the description of the research process carried out in data acquisition and management. This part is based on a common agreed assumption that data life cycle cannot be considered independently from research lifecycle. In the second one a selection of questions already submitted in large-scale international surveys are proposed in order to measure commonality and differences in attitudes as well as in institutional policies.



## Bionotes

---

**Daniela Luzi** is researcher of the National Research Council at the Institute of research on populations and social politics. Her interest in Grey Literature started at the Italian national reference centre for SIGLE at the beginning of her career and continued carrying out research on GL databases, electronic information and open archives. She has always attended the International GL conferences and in 2000 she obtained an award for outstanding achievement in the field of grey literature by the Literati Club. Email: [d.luzi@irpps.cnr.it](mailto:d.luzi@irpps.cnr.it)

**Roberta Ruggieri** is librarian at the Senate of the Republic where she is responsible for the supervision of a digitalization project on Senate parliamentary print documents for the 'I to X Legislature'. Her activity in managing digitalization project also includes document addition and classification in the electronic Senate catalogue. From 2004 she has been collaborating with the Institute for research on populations and social policies of the National Research Council (CNR) in research activities related to the field of Grey literature and Institutional repositories. Email: [biblio.irpps@irpps.cnr.it](mailto:biblio.irpps@irpps.cnr.it)

**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 (ETRD) and currently of the PUMA (Publication Management) & MetaPub, a service oriented and user focused infrastructure for institutional and thematic Open Access repositories looking at the DRIVER vision, <http://puma.isti.cnr.it>. She has coauthored a number of publications dealing with digital libraries. Her activities include integration of grey literature into library collections and web access to the library's digital resources, including electronic journals and databases. She is a member of GreyNet since 2005. Email: [stefania.biagioni@isti.cnr.it](mailto:stefania.biagioni@isti.cnr.it)

## Data sharing in environmental sciences: A survey of CNR researchers

Daniela Luzi\*, Roberta Ruggieri#,  
Stefania Biagioni°, Elisabetta Schiano §

\*CNR-IRPPS, #Senate of the Italian Republic,  
° CNR-IRPPS, § CNR- ISMAR

---

---

---

---

---

---

---

---

### Objectives

#### Analysis of researchers' attitudes in data sharing

- Are researchers willing to share data?
- At which conditions?
- What are their concerns?

#### Analysis of practices and research context

- How research activities are carried out?
- How research data are managed, diffused, preserved and re-used?

Capture perceptions, identify barriers and possible enablers to support open data sharing and contribute to the development of suitable information resources

**SOURCE OF ANALYSIS:**  
CNR Institutes belonging to the  
Department of Earth and Environment  
• data-intensive community  
• multidisciplinary

---

---

---

---

---

---

---

---

### Methods

- Survey design
  - ✓ Analysis of previous international surveys
  - ✓ Identification of the survey target group
  - ✓ Identification of ad hoc questions for CNR
  - ✓ Development of the questionnaire (40 questions)
- On-line questionnaire submission via e-mail
  - ✓ Survey period: 20 June – 15 September
  - ✓ Collection of anonymous responses
- Data analysis

---

---

---

---

---

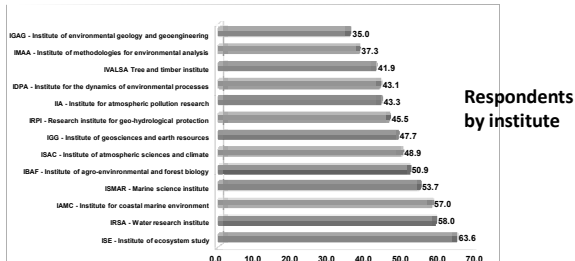
---

---

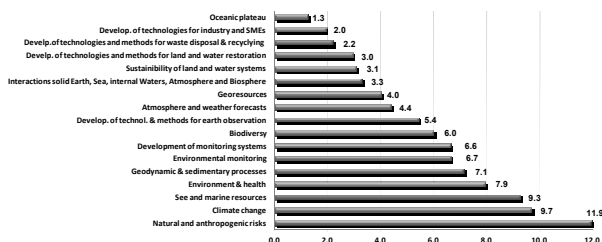
---

## The sample

1087 questionnaires sent  
523 received 48%



## Research context



- Research is mainly funded by national public funds = 47.8% and EU projects = 28.9%
- Research activities are mainly carried out in medium size groups (3 to 7 persons) = 47.7%
- 42% of researchers always work in multidisciplinary groups in International projects, and 37% with other colleagues of the same Institute

## Data acquisition

Mainly experimental (77%)

	Biological	Chemical	Physical	Geological
Land	18.2	30.8	31.9	44.2
Sea	26.4	29.8	30.6	25.8
Internal waters	22.4	36.3	30.0	26.8
Atmosphere	6.5	27.3	42.1	13.4
Biosphere	28.3	24.3	21.0	16.4

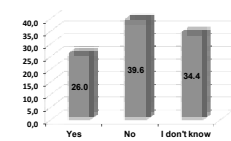
21.2% of respondents used also demographic data

### How data are acquired:

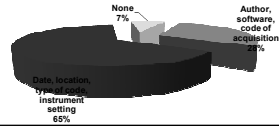
- ✓ Measurements are mainly taken directly by researchers = 53%
- ✓ Data are mainly acquired both in laboratory and in the field = 53%
- ✓ Using instrumentation directly managed by CNR = 83%, but also by other national organizations (26%)

## Data management

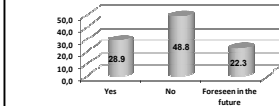
Does your community of reference use standards to manage data?



What type of additional information do you associate with data collected/analysed by yourself?



Does your institute have set up procedures for data preservation?



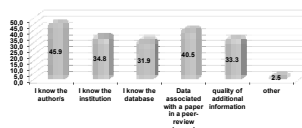
Is anyone in your Institute specifically trained to preserve data?



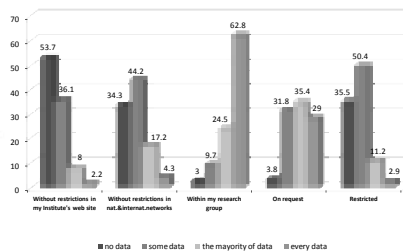
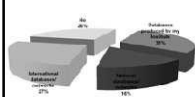
## Data use

• 59% uses data produced by others, mainly in his/her disciplinary field (= 43%)

Data produced by others are considered reliable when..



## Data availability



## In your opinion for which reasons is it important to make research data available and preserve them?

	Very important	Important	Not very important	Not important of all	Missing
The availability of data enhances the transparency of research results	53.9	40.7	3.8	0.6	1.0
When research is publicly funded, data should be available to anyone/ everybody.	50.7	38.6	7.5	2.1	1.1
The availability of data fosters the progress of science (new research is based on pre-existing knowledge)	56.8	38.2	3.6	0.4	1.0
It is a means to validate the results obtained	40.2	43.6	12.6	1.1	2.5
Existing results can be re examined	34.6	46.3	14.5	2.7	1.9
It can promote collaboration among different fields	39.6	45.5	12.8	0.8	1.3
It has a a potential economic value	19.1	39.8	34.2	4.2	2.7
Research data are unique	19.9	40.0	26.8	9.8	3.6
The availability of data reduces the duplication of research efforts	36.9	35.9	18.9	6.3	1.9

### In your opinion what are the main obstacles of data sharing?

	Very important	Important	Not very important	Not important of all	Missing
Lack of funds	31.4	30.6	27.9	5.4	4.8
Lack of standards	25.8	46.3	18.9	3.1	5.9
It requires too much time	16.1	38.0	32.7	7.8	5.4
Difficulties in adoption of standard	13.0	38.4	33.7	8.4	6.5
No technical support	31.4	41.9	16.4	3.6	6.7
There are no archives to submit to	23.3	37.3	23.9	9.2	6.3
Procedures of data sharing are too complicated	10.7	33.8	38.4	10.7	6.3
Loss of data control	19.9	31.4	30.4	12.6	5.7
Data may be used and/or misinterpreted	22.8	35.6	25.4	10.5	5.7
Data are not evaluated like papers in scientific journals	37.5	31.5	20.8	5.0	5.2
Loss of exclusivity/originality of the work	26.4	29.4	30.0	8.6	5.5

---

---

---

---

---

---

---

---

---

---

### What conditions would you require to submit your research data to an open archive?

	Very important	Important	Not very important	Not important of all	Missing
I will be able to update data after submission	60.2	30.8	4.2	1.7	3.1
I will be able to delete data	31.2	33.7	22.0	7.3	5.9
I know who is using data, when and for which purpose	53.5	27.5	11.7	3.6	3.6
Be contacted if someone wants to use my data	52.0	30.2	12.0	2.3	3.4
Receive a formal acknowledgment	35.4	36.1	20.3	4.0	4.2
Be reassured about long-term data preservation	38.6	39.8	13.4	3.3	5.0
Simple procedures to deposit data	52.6	37.1	5.5	0.6	4.2
Receive additional funds	24.7	39.2	27.5	4.4	4.2
Receive the same evaluation received for publications	41.1	37.5	13.0	3.8	4.6

---

---

---

---

---

---

---

---

---

---

### Conclusions

#### High level of awareness

- 48% of voluntary responses to the questionnaire;
- Data sharing and preservation are very important as it fosters the progress of science and enhances the transparency of research
- Publicly funded research should be available to anyone

#### Willingness to share correlated with actions

- 2.2% are not, 69.8% depending on the circumstances
- The majority of researchers store their data in local, national and international databases, but they tend to share only a fraction of them

Obstacles are connected with data management and Concerns are mainly related to the formal acknowledgment of time consuming data sharing

Motivations: a friendly environment to deposit data related to data storage, management and re-use

---

---

---

---

---

---

---

---

---

---

## Tracking the Influence of Grey Literature in Public Policy Contexts: The Necessity and Benefit of Interdisciplinary Research

Bertrum H. MacDonald, Elizabeth M. De Santo,  
Kevin Quigley, Suzuette S. Soomai, and Peter G. Wells  
Dalhousie University, Halifax, Nova Scotia, Canada

The pivotal role that scientific information (much of it grey literature) can play in the search for solutions to serious global environmental problems is receiving growing attention by a diversity of researchers recently. A.P.J. Mol, professor of environmental policy at Wageningen University, for example, stated that “it is the production, the processing, the use and the flow of, as well as the access to and the control over, information that is increasingly becoming vital in environmental governance practices.... and the motivations and sources for changing unsustainable behaviour are increasingly informational” (2010).

How information functions within the interface between science and policy is only weakly understood, in part because most studies have been conducted through single disciplinary lenses. Moreover, determining the life cycles of scientific information and developing an understanding of the use and influence of this information are not trivial tasks. We believe that an appreciable increase in understanding can be achieved through an interdisciplinary perspective and a comparative approach employing a suite of research methodologies to document information pathways. In particular, we contend that interdisciplinary research, drawing on “information science and management,” “marine environmental science,” “marine policy development,” “fisheries science and management,” and “public policy,” can substantially increase understanding of the processes by which scientific information is incorporated into environmental policy decisions (Figure 1). This innovative, evolving interdisciplinary perspective enables addressing the question “what role and influence does grey literature have in marine environmental policy and decision-making processes” in an informative, holistic manner not feasible otherwise.

Environmental problems and related policy decisions are multidimensional, as the *State of the Planet Declaration* issued at the conclusion of the recent Planet Under Pressure: New Knowledge Towards Solution conference (London, March 2012) emphasized. “In one lifetime,” the declaration stated, “our increasingly interconnected and interdependent economic, social, cultural, and political systems have come to place pressures on the environment that cause fundamental changes in the Earth system....But the same interconnectedness provides the potential for solution: new ideas form and spread quickly, creating the momentum for the major transformation required for a truly sustainable planet.” As this paper will show, multidimensional analysis provided by an interdisciplinary perspective is essential for understanding the role of scientific information in the science-policy interface in marine environmental decision-making.

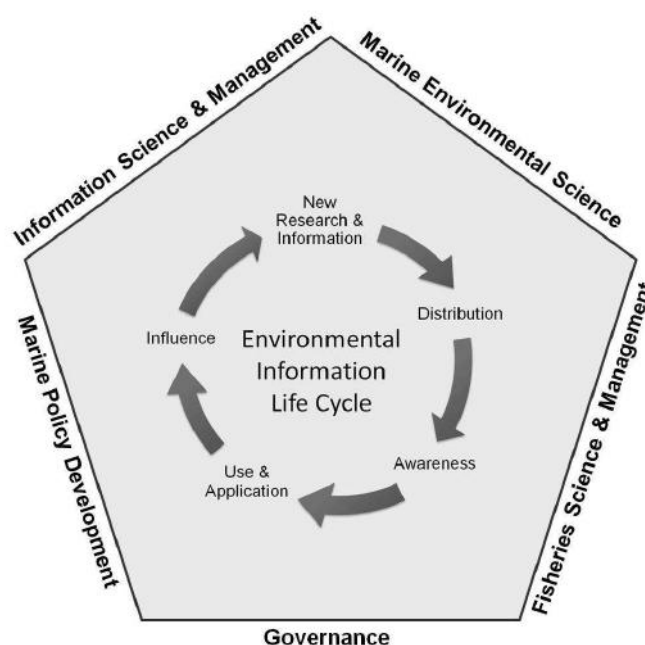


Figure 1 - Interdisciplinary Analytical Framework

## 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](http://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](mailto:bertrum.macdonald@dal.ca)

**Elizabeth M. De Santo** is an Assistant Professor in the Marine Affairs Program at Dalhousie University, Halifax, Canada. Dr. De Santo holds a BA (Honors, Zoology) from Connecticut College, a Master of Environmental Management from Duke University, a Master of Science (International Relations) from the London School of Economics and Political Science, and a MPhil & PhD (Geography & Law) from University College London. Prior to her appointment at Dalhousie University, Dr. De Santo was the Marine Protected Areas Coordinator with the International Union for Conservation of Nature, based in Washington, DC; Program Manager with the World Environment Center, New York; and a Researcher with the American Museum of Natural History, New York. Email: [elizabeth.de.santo@dal.ca](mailto:elizabeth.de.santo@dal.ca)

**Kevin Quigley** is Associate Professor and Director of the School of Public Administration at Dalhousie University, Halifax, Canada. Dr. Quigley holds a BA (English) from Queen's University (Kingston, Canada), a MSc (Public Administration and Public Policy) from the London School of Economics and Political Science, and a PhD in Public Policy Studies from Queen's University Belfast. Prior to his appointment at Dalhousie University he was a Post-Doctoral Fellow at the University of Edinburgh and a Visiting Scholar at the American Political Science Association in Washington DC. He has also held public sector appointments with the Government of Ontario, Canada. His research focusses on critical infrastructure protection, security, risk regulation, and public policy. His research has been funded by the Social Sciences and Humanities Research Council of Canada, the Canada School of Public Service, Defence Research Development Canada, Public Safety Canada and the UK's Economic and Social Research Council. Email: [Kevin.Quigley@dal.ca](mailto:Kevin.Quigley@dal.ca)

**Suzette 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, Suzette 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: [suzette.soomai@dal.ca](mailto:suzette.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](mailto:oceans2@ns.sympatico.ca)



I search



I find



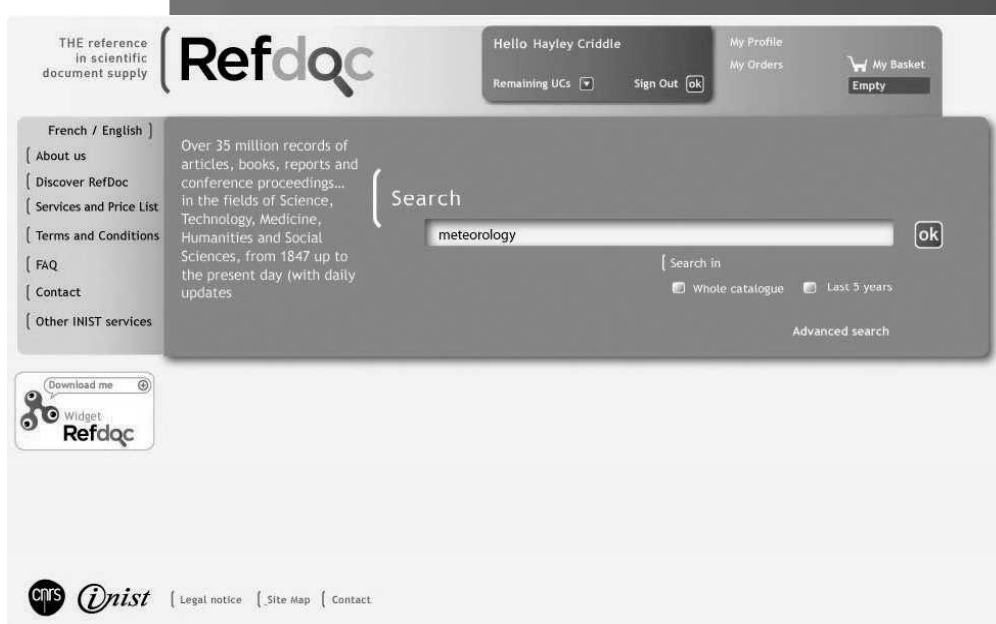
I order

January 2010 – all scientific  
and technical information  
available in just 3 CLICKS

Over 35 million records  
of articles, books,  
reports and conference  
proceedings from 1847  
up to the present

( Refdoc.fr )

DISCOVER ALL THE FEATURES AND FUNCTIONS AVAILABLE AT [www.refdoc.fr](http://www.refdoc.fr)





## Grey communities

### An empirical study on databases and repositories

**Hélène Prost**, Institute for Scientific and Technical Information, INIST-CNRS, France

**Joachim Schöpfel**, Charles de Gaulle University Lille 3, France

**Problem/goal** - In 2012, the GreyNet network celebrates its 20th anniversary, with nearly 300 contributors from 30 different countries. In recent years, GreyNet directed its activities towards open access through the launch of OpenGrey, and the creation of the GreyNet LinkedIn group marked its entry into social networks.

However, does the network attain and aggregate all scientists, academics and information professionals interested in the field of grey literature and contributing to its knowledge? Which is the potential for development of GreyNet?

Our paper will explore the grey communities outside Grey Literature Network Service. It will try to discover potential members for GreyNet. Our study is following two other papers on scientific and professional participants of GreyNet, the first based on a citation analysis of the first five conferences on grey literature (Schöpfel et al., 2005), the second describing the introduction of the new OpenGrey (formerly OpenSIGLE) service (Farace et al., 2009).

**Research method/procedure** - The GreyNet community is defined by the authors who publish in the GreyNet newsletter, in The Grey Journal, or in the proceedings of the annual conferences on grey literature. In order to detect potential contributors outside the GreyNet community, we shall undertake a bibliographic research on grey literature as a topic of publication, e.g. we shall search for original papers, articles, reports, theses and so on on grey literature, discarding the cited publications.

The search for references will focus on the following bibliographic and scientometric databases: Web of Science (Thomson), Scopus (Elsevier), Library Information Science & Technology Abstracts (EBSCO), Library and Information Science Abstracts (CSA), Pascal and Francis (INIST-CNRS). It will also include the E-LIS repository with e-prints in library and information science and, if suitable, a search in the OALster database via the OCLC WorldCat. The objective is to get an exhaustive list of the scientific production on grey literature published between 2000 and early 2012.

The obtained information should allow identifying authors, their institution and country, together with preferred journals, collections and so on where they publish. We shall present a general description of this corpus and then identify the major authors. Some of them frequently use the concept of “grey literature” while others publish on topics related to grey literature, especially on dissertations and theses (citation analysis...), without this concept.

**Anticipated results** By comparison of this corpus with the GreyNet directory, our study is expected to produce a list of authors - scientists, scholars and professionals - who may be interested to join the GreyNet community, via the GL conferences, the subscription to the newsletter and/or journal, or via the social network on LinkedIn. This list should be helpful for the further development of GreyNet.

## Bionotes

**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](mailto:helene.prost@inist.fr)

**Joachim Schöpfel** is Head of the Department of Information and Library Sciences at the Charles de Gaulle University of Lille 3 and Researcher at the GERiiCO laboratory. He is interested in scientific information, academic publishing, open repositories, GL and usage statistics. He is a member of GreyNet and euroCRIS. He is also the Director of the National Digitization Centre for PhD Theses (ANRT) in Lille, France. [joachim.schopfel@univ-lille3.fr](mailto:joachim.schopfel@univ-lille3.fr)

## Use Pattern of Archives on the History Of Mysore

**Dr. N. Chowdappa**, BMS College of Engineering, Bangalore, India

**L. Usha Devi**, Bangalore University, Bangalore, India

**Dr. C.P. Ramasesh**, University of Mysore, Mysore, India

Records on the Administration of the Princely State of Mysore and Mysore History form rare collections for historians who venture to study the history of Mysore State under the rule of the Wadiyars, the rulers of Mysore State. These rare materials and archives have been carefully preserved at the Archival Section of the University of Mysore and also at other libraries in Mysore city. The present study furnishes the type of archival materials available at the University Library, Oriental Research Institute and the Karnataka State Archives, Mysore Division. Further, the study depicts the purposes of accessing archives and the use pattern of these rare archival sources on Mysore history by the research scholars, students and teachers in the discipline of Karnataka/Mysore History and allied fields. The present study also projects the extent of dependency of scholars from various professional fields, for information sources on Mysore History. The study also projects the rare collections of manuscripts and books of Tipu's Library.

### Bionotes

**Narayanappa Chowdappa** obtained a Master's Degree in Geology and also Master's Degree in Library and Information Science from Bangalore University, Bangalore with a distinction. Obtained Doctorate degree for his Thesis on "Organization and Use Patterns of Grey Literature in Engineering Research Institutions" from the University of Mysore, Mysore in Library and Information Science. Serving as Chief Librarian at BMS College of Engineering, Bangalore for the last 28 years. Special interest includes promoting the use of grey literature among faculty and researchers in engineering discipline. Organized 24 professional and extension programs for Teachers and Librarians in Science and Technology. Served as resource person in library and information science, and Academic Counselor for Indira Gandhi National Open University, New Delhi. Presently Dr. Chowdappa is holding the position of President, AKELPA - All Karnataka Engineering College Library Professionals Association, Bangalore. Areas of professional interest are: Scholarly Communications, Facilitating Research Programs and Reference Service.

Email: [ncbmsce@yahoo.co.in](mailto:ncbmsce@yahoo.co.in)

**L. Usha Devi** obtained her Bachelor's Degree in Science and Master's Degree in Library and Information science from the University of Mysore, Mysore. She has been working as an Assistant University Librarian at Bangalore University, Bangalore for the last 30 years. Smt Ushadevi also possess M.Phil Degree in Library and information Science and Post graduate diploma in English language teaching, Postgraduate Diploma in Human Resource Management also. She has 5 years of postgraduate teaching experience in Library and information science and presented and published 14 research papers at National and International level seminars and conferences. She is a life member of professional associations: ILA, IASLIC, IATLIS, KALA. and AKELPA. Her areas of research interest are: User studies, Organizations and Use of Grey literature, Citation Studies. Email: [ushachowdappa@yahoo.in](mailto:ushachowdappa@yahoo.in)

**Chowbiny Ramasesh** obtained his Master's Degree in Philosophy from the University of Mysore in 1976 with specialization in Vedanta Philosophy. He also obtained a Master's Degree (1978) and Doctorate Degree (1989) in Library and Information Science from the University of Mysore. Ramasesh is the recipient of Dr. S.R. Ranganathan Memorial Gold Medal for securing first rank in the Master's Degree. He served as professional librarian for three decades and is responsible for the organization of several extension programmes. Worked as the Deputy Director of Centre for Information Science and Technology (CIST), Mysore and coordinated for the implementation of Quality Procedures under ISO 9001 Standards of Quality Management System. Compiled Quality Manual, delivered special lectures and served as Quality Auditor/ Performance Auditor at CIST. Presently serving as University Librarian of the University of Mysore and supervising research programs in the field of 1) Grey Literature 2) Institutional Repositories of Heritage Collection and 3) Use Pattern of Online Journals. Email: [cpramasesh@gmail.com](mailto:cpramasesh@gmail.com)

## USE PATTERN OF ARCHIVES ON HISTORY OF MYSORE

N. Chowdappa, L. Usha Devi and C.P.Ramasesh




---

---

---

---

---

---

---

## HISTORY OF MYSORE (Subject Coverage)

1. Dynasty of Wadiyars
2. Rule under Hyder Ali and Tipu Sultan
3. Mysore Wars
4. Administration under Wadiyars
5. Culture and Heritage of Mysore
6. Education and Literature
7. Art, Architecture and Music




---

---

---

---

---

---

---

## DOCUMENTS OF HISTORICAL IMPORTANCE

1. Manuscripts
2. Collection of Letters
3. Maps and Charts
4. Monographs
5. Books
6. Directories
7. Handbooks
8. Reports
9. Serials




---

---

---

---

---

---

---

Important Library Holdings (I)		
SL. No.	LIBRARIES IN MYSORE	Total Collection
01	Tipu's Library Collection (TLC)	2715
02	Mysore Palace Library (MPL)	16550
03	Saraswathi Bhandar (SBM)	4120
04	Department of Archaeology and Museum (DAM)	5620




---

---

---

---

---

---

---

---

Sl. No.	Important Library Holdings (II)	Total Collection
05	Mysore University Library (MUL)	621550
06	Oriental Research Institute (ORI)	112180
07	State Archives, Mysore (SAM)	38418




---

---

---

---

---

---

---

---

RARE COLLECTIONS IN TIPU'S LIBRARY : 1799 A.D. (Languages : Persian, Arabic, Turkish, Hindi and Kannada)		
Sl. No.	Subject Areas	Number
01	Philosophy and Religion	239
02	Koran and Commentaries	120
03	Linguistics	74
04	Literature	237
05	Culture and Tradition	46
06	Collection of Letters	53
07	Jurisprudence	95
08	Science and Technology	108
09	Astronomy	22
10	History	118
	<b>Total Rare Works</b>	<b>1112</b>
	<b>Total Library Collection</b>	<b>2715</b>

---

---

---

---

---

---

---

---

HOLDINGS ON HISTORY OF MYSORE (In the Libraries Selected for the Study)		
Sl. No.	Selected Libraries in Mysore	Archival Collection
01	Mysore University Library (MUL)	8095 (1.3%)
02	Oriental Research Institute (ORI)	3018 (2.7%)
03	State Archives, Mysore (SAM)	16506 (43.0%)
	TOTAL COLLECTION	27619 (3.6%)

---

---

---

---

---

---

---

---

PURPOSES OF THE USE OF ARCHIVES				
Sl. No.	Purposes of the Use Of Archives	Percentage of Use		
		MUL	ORI	SAM
01	Teaching	14%	12%	22%
02	Research	21%	23%	34%
03	Student Projects	48%	41%	13%
04	Publish News Items	17%	24%	31%

---

---

---

---

---

---

---

---

USE OF ARCHIVES AT MUL				
Sl. No.	Category of Users	Year-wise Visitors		
		2009	2010	2011
01	Teachers	409	419	426
02	Research Scholars	601	586	616
03	Students	1410	1316	1418
04	Public and Journalists	399	447	489
	Total	2819	2768	2949

---

---

---

---

---

---

---

---

### USE OF ARCHIVES AT ORI

Sl. No.	Category of Users	Year-wise Visitors		
		2009	2010	2011
01	Teachers	101	112	118
02	Research Scholars	196	210	228
03	Students	356	359	413
04	Public and Journalists	211	247	243
	<b>Total</b>	<b>864</b>	<b>928</b>	<b>1002</b>

---

---

---

---

---

---

---

---

### USE OF ARCHIVES AT SAM

Sl. No.	Category of Users	Year-wise Visitors		
		2009	2010	2011
01	Teachers	816	793	803
02	Research Scholars	1212	1256	1229
03	Students	427	439	471
04	Public and Journalists	805	917	1107
	<b>Total</b>	<b>3260</b>	<b>3405</b>	<b>3610</b>

---

---

---

---

---

---

---

---

### SUMMARY OF FINDINGS

1. Archives are being extensively used for major research programmes and projects of Master's Degree students.
2. Study revealed increasing trend in use of the archives.
3. Archives on thrust areas, especially the records on Wadiyars available at SAM Library is being frequently used.
4. Archives of utmost importance are related to the areas of
  1. *Wadiyars of Mysore : Life and Contribution ;*
  2. *Education and Cultural Heritage of Mysore*
5. Rare manuscripts and books of historical importance have to be protected using preservation methods at MUL and ORI.
6. Digitisation of materials of historical importance has to be completed on priority at MUL and ORI.

---

---

---

---

---

---

---

---

## Data Analytics: The next big thing in information

**June Crowe and J.R. Candlish**

Information International Associates, Inc. (IIa), United States

Information is now available in an overabundance, so much so, that distinguishing the noise from the signal has become very problematic. In the past, the collection and storage of information was the primary issue. In addition, the overabundance of data has made relational databases become overloaded and unable to provide the needed timely analysis required by some organizations. However, some databases, such as Factiva, are now offering Beta versions of a more analytic nature.

Currently, we have massive amounts of data both structured and unstructured, that need to be analyzed in an interactive, as well as in a time sensitive manner. Data analytical tools have emerged to solve this problem. Some of the names in the field include: R Studio, SAP Hana, Platfora, Aster Dater, Green Plum, Teradata, Netezza, ParAccel, Vertica, Hadoop, MapReduce, Insight, Bitext, Attensity, and Lexalytics. Grey Literature repositories, libraries, and information centers are well positioned to take advantage of these new tools. The current trend is to make grey literature more easily discoverable, accessible, and with the new data analytical tools, more easily analyzed.

This study proposes to survey the membership of the Grey Literature community on what data analytic tools may already be in use and what is planned for the near future. The survey will include questions that inquire the motives/demands for using these tools, what tools are being used or contemplated being used and why, as well as how the outcome will assist in making grey literature more visible.

### 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. She has extensive experience in the management and operations of library services across government, public, academic, and special libraries. At IIa she manages the open source research division which focuses on medical, socio-cultural, science and technology and business research. Her primary interests are open source information in Grey Literature, digital repositories, and open source intelligence tools. Email: [jcrowe@iiaweb.com](mailto:jcrowe@iiaweb.com)

**J.R. Candlish** is a Technical Assistant and GIS Analyst at Information International Associates, Inc. (IIa). He received his M.S. in Biosystems Engineering Technology from the University of Tennessee, Knoxville and his B.S. in Environmental Studies: Natural Resources from Sewanee: The University of the South. At IIa he provides secretariat services to CENDI agencies, the federal scientific information manager's group and also fulfills a mapping role for the USGS National Biological Information Infrastructure's Southern Appalachian Information Node (SAIN). [jcandlish@iiaweb.com](mailto:jcandlish@iiaweb.com)

**Data Analytics: The Next Big Thing in Information**

Rome, Italy  
November, 2012

June Crowe and J.R. Candlish, Information International Associates, Inc., Oak Ridge, TN.

Iia Proprietary Information

---

---

---

---

---

---

---

---


**Abstract**

**Data Analytics: The Next Big Thing in Information**

Information is now available in an overabundance, so much so, that distinguishing the noise from the signal has become very problematic. In the past, the collection and storage of information was the primary issue. Currently, we have massive amounts of data both structured and unstructured, that need to be analyzed in an iterative, as well as in a time sensitive manner. In the meantime, data analytical tools have emerged to solve this problem.

Grey literature repositories, libraries, and information centers are well positioned to take advantage of these new tools. The current trend is to make grey literature more easily discoverable, accessible, and with the new data analytical tools, more easily analyzed.

We created and administered a survey to the Grey Literature Community to get a snapshot of the Community's use, planned use, and knowledge of Big Data as it affects grey literature. The results are reported in these slides.

 Iia Proprietary Information 11/10/2012 2

---

---

---

---


---

---


---

---

**Survey Participants from Around the World**



Total No.: 48

 Iia Proprietary Information 11/10/2012 3

---

---

---

---

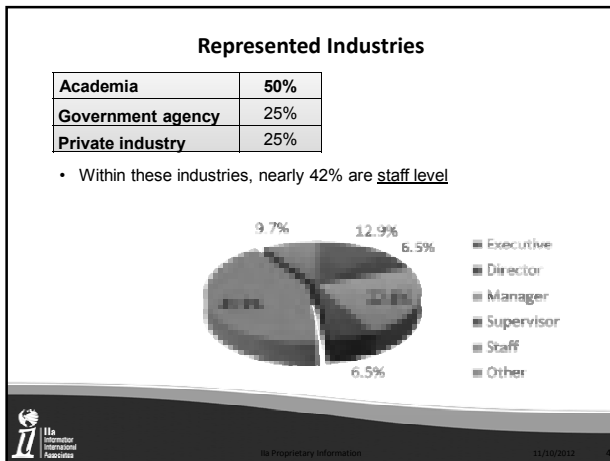
---

---

---

---






---

---

---

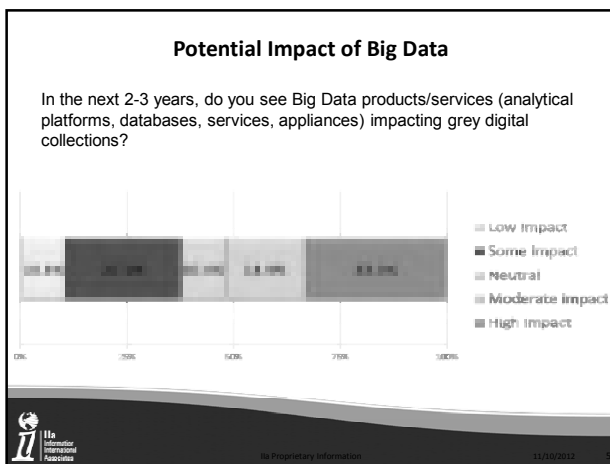
---

---

---

---

---




---

---

---

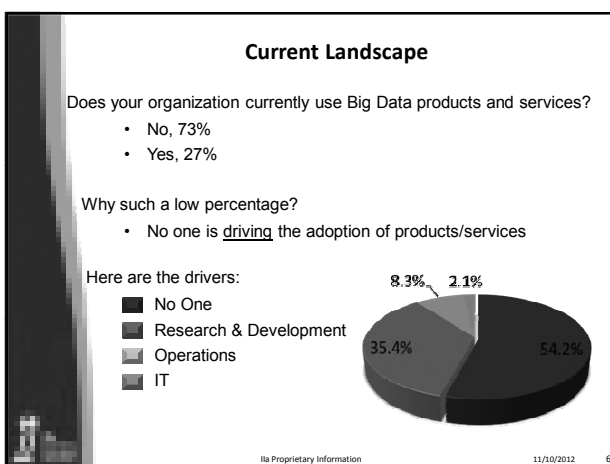
---

---

---

---

---




---

---

---

---

---

---

---

---

### Current Landscape

Indicate your level of expertise in Big Data: **Novice, 74.4%**

Within the breadth of product/services currently available, the three most recognized were:

1. SAP
2. SAS
3. Google BigQuery

Interesting Finding: 81.4% of survey takers have not seen any Big Data products/services demonstrated.

However: current/planned analytic Big Data projects favored Web analytics (30.4%), Predictive analytics (25.6%), and Real-time analytics (15.4%) respectively.

IIa Proprietary Information

---

---

---

---

---

---

---

---

### Importance of Big Data

According to a survey The Pew Research Center undertook about Big Data, most respondents think that the rise of Big Data is a huge positive for society. Do you think that Big Data will be a huge positive for your organization?

<b>Definitely (value of 5)</b>	<b>30.8%</b>
<b>Moderate (4)</b>	<b>25.6%</b>

**>56% of responses!**  
Only 10% neutral

Regarding Mobile Devices: A strong percentage (36.1%) of survey takers feel it is only "Somewhat Important" to access Big Data Results via mobile devices.

IIa Proprietary Information 11/10/2012 8

---

---

---

---

---

---

---

---

### Barriers to Adopting Big Data Products/Services

Please rank the following potential barriers within your organization that may impede the adoption of Big Data products/services. Please rank each on a scale of 1-5 with 1 being the least and 5 the greatest.

	1	2	3	4	5
<b>Lack of Personnel</b>	11.1%	11.1%	19.4%	25.0%	<b>33.3%</b>
<b>Lack of Resources</b>	8.6%	11.4%	14.3%	20.0%	<b>45.7%</b>

IIa Proprietary Information 11/10/2012 9

---

---

---

---

---

---

---

---

### Goals to Utilize Big Data Products/Services

If it were possible to select any of the Big Data products/services, which of the following common data goals would be most relevant in your situation? Please rank each on a scale of 1-5 with 1 being the least and 5 the greatest.

	1	2	3	4	5
<b>Data visualization</b>	2.8%	8.3%	19.4%	30.6%	<b>38.9%</b>
<b>Data discovery</b>	5.6%	8.3%	8.3%	30.6%	<b>47.2%</b>
<b>Data mining analytics</b>	2.8%	2.8%	27.8%	22.2%	<b>44.4%</b>



Ila Proprietary Information

11/10/2012 10

---

---

---

---

---

---

---

---

### Potential Concerns and Local Economy

As Big Data capabilities evolve, do you think that biased reporting will be a cause of concern across multiple facets of applicability (economic, political, social, scientific, health, etc.)?

Least Concern	Somewhat Concerned	<b>Moderate Concern</b>	Moderately High Concern	High Concern
2.9%	2.9%	<b>52.9%</b>	26.5%	14.7%

In your opinion, will Big Data initiatives benefit your local economy?

Least	Somewhat	<b>Moderate</b>	Moderately High	Most
5.7%	25.7%	<b>42.9%</b>	20.0%	5.7%



Ila Proprietary Information

11/10/2012 11

---

---

---

---

---

---

---

---

### Metrics on Survey Participants

Total Number of Survey Takers: **48**

Number of Completed Surveys: **32 (66.7%)**

#### Trend in Responses

- Of the 24 questions, the first third of the questionnaire had the most participation.
- The second third had an average of 36-39 participants
- The last third had a range of responses from 19-36.

11/10/2012

Ila Proprietary Information

12

---

---

---

---

---

---

---

---

## GreyNet's Enhanced Publications Project: Tracking and Backtracking Data

**Dominic Farace and Jerry Frantzen**, GreyNet International, Netherlands

**Christiane Stock**, INIST-CNRS, France

**Paula Witkamp**, Data Archiving and Networked Services, DANS, Netherlands

In 2011, GreyNet presented the first results of a two year project on Enhanced Publications (EPP). This year, the project focusses on the acquisition of research data, their cross-linking to existing full-text documents, and the establishment of a workflow for future publications. Enhanced publications combine and link research data to full-texts, other supplementary materials, as well as post-publication data.

Results in the first leg of the project indicate that sixty percent of the surveyed authors base their research on empirical and/or statistical data. And, two-thirds of those authors remarked that their data are still available for archiving purposes. These respondents also express a willingness to share their data and hold to the opinion that both the data producer as well as the prospective user would stand to benefit.

This second leg of the project rests on the approach taken in facilitating the acquisition process. The primary instrument used will be the OpenGrey Repository that houses GreyNet's collection of conference preprints. By backtracking to the existing metadata records in OpenGrey and by communicating directly with the authors of those records, another way will be opened for further cooperation between data producer and data provider. In addition, the subsequent cross-linking between OpenGrey and the DANS EASY Repository, where GreyNet's research data will be stored, stands to better serve the needs of the grey literature community in which open access to research data is a prerequisite.

There are no direct costs associated with this project. Each of the partners is committed to allocate human and material resources needed to carry out their related tasks.

### Project Deliverables 2011

Linking full-text grey literature to underlying research and post-publication data (Full-Text)  
[http://www.greynet.org/images/GL13-S4P\\_Farace\\_et\\_al.pdf](http://www.greynet.org/images/GL13-S4P_Farace_et_al.pdf)

Linking full-text grey literature to underlying research and post-publication data (PowerPoint)  
[http://www.greynet.org/images/GL13-S4S\\_Farace\\_et\\_al.pdf](http://www.greynet.org/images/GL13-S4S_Farace_et_al.pdf)

Shining a light on grey literature (PowerPoint)  
[http://www.greynet.org/images/GL13-PSS\\_Edwards\\_et\\_al.pdf](http://www.greynet.org/images/GL13-PSS_Edwards_et_al.pdf)

## Bionotes

**Dominic J. Farace** is Director of TextRelease, an Amsterdam based information bureau specializing in grey literature and networked information. He is a native Louisianan and holds two degrees in sociology from Creighton University (BA) and the University of New Orleans (MA). His doctoral dissertation in social sciences is from the University of Utrecht, The Netherlands, where he has lived and worked for the past twenty-seven years. After six years heading the Department of Documentary Information at the Royal Netherlands Academy of Arts and Sciences (Swidoc/KNAW), he founded GreyNet, Grey Literature Network Service, in 1993 and has since been responsible for the international GL-Conference Series. In this capacity, he serves as Program and Conference Director as well as managing editor of the conference proceedings. Since 2004, he is a Guest Lecturer on Grey Literature in the Masters Program at the University of Amsterdam; Instructor of Grey Literature via UNO Distance Education, and Editor of TGJ, The Grey Journal. Email: [dominic.farace@textrelease.com](mailto:dominic.farace@textrelease.com)






**Jerry Frantzen** graduated in 1999 from the College of Amsterdam in Library and Information Science. Frantzen is the technical editor of The Grey Journal (TGJ). And, since 1996, he is affiliated with GreyNet, Grey Literature Network Service, as a freelance technical consultant. Email: [info@greynet.org](mailto:info@greynet.org)

**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 OpenSIGLE. 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](mailto:christiane.stock@inist.fr)

GL14  
Fourteenth International Conference on Grey Literature

**GreyNet's Enhanced Publications Project:  
Tracking and Backtracking Data**

Dominic Farace and Jerry Frantzen, GreyNet International, Netherlands  
Christiane Stock, INIST-CNRS, France  
Paula Witkamp, Data Archiving and Networked Services, DANS, Netherlands

TRACKING INNOVATION THROUGH GREY LITERATURE  
National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

GL14  
Fourteenth International Conference on Grey Literature

**Enhanced Publication Project – EPP  
Contents**

- Data Acquisition Phase
- Problems Acquiring Data
- Cross-Linking Approach
- Data Entry Guidelines
- Sample Record in DANS Data Archive
- Sample Record in OpenGrey Repository
- Tracking and Backtracking Data
- Community Awareness
- Who Stands to Benefit?

TRACKING INNOVATION THROUGH GREY LITERATURE  
National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

GL14  
Fourteenth International Conference on Grey Literature

**Enhanced Publication Project – EPP  
Data Acquisition Phase**

- Project Summary - Part One
- Project Goals - Part Two

The definition of an **enhanced publication** is borrowed from the DRIVER-II project, “a publication that is enhanced with three categories of information: 1) research data, 2) extra materials, and 3) post-publication data”. <http://www.driver-repository.eu>

TRACKING INNOVATION THROUGH GREY LITERATURE  
National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

GL14 Fourteenth International Conference on Grey Literature

**Enhanced Publication Project – EPP**  
**Problems Acquiring Data**

- Open versus Restricted Data Access
- Non-English Language Variable Headings
- Description of data different from FT publication
- Centralized versus Decentralized data input
- Data entry in a cross-boader archive

TRACKING INNOVATION THROUGH GREY LITERATURE  
National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

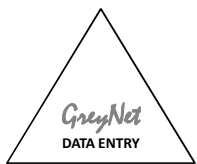
---

---

GL14 Fourteenth International Conference on Grey Literature

**Enhanced Publication Project – EPP**  
**Cross-Linking Approach**

OpenGrey Repository



DANS Data Archive Other Data Archives

TRACKING INNOVATION THROUGH GREY LITERATURE  
National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

GL14 Fourteenth International Conference on Grey Literature

**Enhanced Publication Project – EPP**  
**Data Entry Guidelines**

- Preferred and Accepted File Formats
- Depositing Data in 7 EASY Steps

TRACKING INNOVATION THROUGH GREY LITERATURE  
National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---



Fourteenth International Conference on Grey Literature

**Enhanced Publication Project – EPP**

**Community Awareness**

Enhanced publication inherently **contributes to the review process** of grey literature as well as the **replication of research** and improved **visibility of research** results in the **scholarly communication** chain.

TRACKING INNOVATION THROUGH GREY LITERATURE  
 National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

Fourteenth International Conference on Grey Literature

**Enhanced Publication Project – EPP**

**Who Stands to Benefit?**

- Researchers and contributing Authors
- GreyNet's Collection
- International Grey Literature Community
- Grey Literature as a Field in Library and Information Studies
- Public Understanding of Grey Literature

TRACKING INNOVATION THROUGH GREY LITERATURE  
 National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

Fourteenth International Conference on Grey Literature

**Enhanced Publication Project – EPP**

*Data Sharing begins with Data Publication*

Your contribution will make the difference!

**Thanks!**

TRACKING INNOVATION THROUGH GREY LITERATURE  
 National Research Council, CNR Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---



## Collection of Conference Proceedings and Improving Access to the Full Text of Proceedings

**Misa Hayakawa, Shun Nagaya, Mayuki Gonda, Takeyasu Fukazawa, Minoru Yonezawa, and Keizo Itabashi,** Japan Atomic Energy Agency, JAEA, Japan

In this presentation, we would like to introduce the present status of the collection of conference proceedings (proceedings) in the JAEA Library and the improvement of access to the full text of proceedings.

Generally, proceedings contain the information on the latest research trends, needs of the proceeding is normally high in the fields of science and technology. However, it is difficult to collect at the libraries.

Proceedings have been published mainly in conventional book form. But, in recent years, their publication form has been changing. The proceedings published on the Internet, accessibility to the full text has been relatively improved. But since most of these proceedings are published in various places such as website of scientific societies, special website of the conference, cross-search of these sites is difficult.

As a good practice for the improvement of accessibility to the full text of proceedings, we would like to introduce the case of the JAEA Library. In the JAEA, researchers are requested to apply before they make presentations of their research. After presentation, it is necessary to submit to the library offprint of the paper. So, the library is a single point of management for the information of presentations performed by researchers. In this system, the authority file for the name of conferences is introduced. Because there are various notations in the same conference.

We disseminate these information via Internet from JOPSS (JAEA Originated Papers Searching System). Through this system, information of conference presentations by the JAEA researchers is available.

In addition, we provide the link to proceedings on the Internet at JOPSS since 2011. That is expected to enhance the accessibility to the full text. However, the link to the full text is not permanent because conference secretariat is often a temporary body. And URL on the Internet frequently changes. We Investigated the present statues of proceedings on the Internet that including JAEA researchers have contributed.

Additionally, we examine how to provide bibliographic information to improve access to proceedings that published in conventional book form. So, we would like to examine the problems of the availability of proceedings as a whole.

### Bionotes

**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). E-mail: [hayakawa.misa@jaea.go.jp](mailto:hayakawa.misa@jaea.go.jp)

**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](mailto:gonda.mayuki@jaea.go.jp)

## Collection of Conference Proceedings and Improving Access to the Full Text of Proceedings

Misa Hayakawa, Shun Nagaya,  
Mayuki Gonda, Takeyasu Fukazawa,  
Minoru Yonezawa, Keizo Itabashi  
*Japan Atomic Energy Agency, Japan*

1

---

---

---

---

---

---

---

---

### 1. Introduction

#### 1.1 Conference Proceedings, a type of "grey literature"

- Conference Proceedings (Proceedings) contain papers that are presented at a conference in a specific field.
- Proceedings are published more quickly than journal papers, and they contain the latest research. Proceedings are important sources for investigating research trends.
- They are often ***distributed only to participants***. It is difficult for libraries to collect them because they are not commercially available in many cases.
- Proceedings used to be published as journal special issues or books. Nowadays, Proceedings are also published as CD-ROMs or USBs, and some are published on the Internet.
- Proceedings on the Internet are published at various locations, such as publishers' or conference websites. Their availability may depend on the sustainability of the websites.

⇒ Proceedings are a type of "grey literature."



2

---

---

---

---

---

---

---

---

### 1. Introduction

#### 1.2 Purpose of this study

- We consider how to improve access to the Proceedings, especially on the Internet.
- The number of Proceedings on the Internet has increased, it is important to provide full text to researchers.
- We would like to introduce the JAEA library as a good option for improving accessibility to the full text of Proceedings.
- We then introduce a new approach to provide conference information and hyperlinks to the full text of Proceedings on the Internet.



3

---

---

---

---

---

---

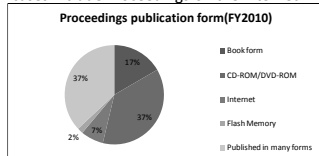
---

---

## 2. Publication forms of Proceedings and the library collection

### 2.1 Survey of the forms of Proceedings publications

- We surveyed the forms of Proceedings publications for the Proceedings written by the JAEA (Japan Atomic Energy Agency) researchers in FY2010.
- We classified the Proceedings by publication form other than those published as a journal or institutional report.
- 17% of the Proceedings are published as books. On the other hand, 37% Proceedings are published on CD-ROM.
- "Published in many forms" includes books, CD-ROM, etc. Many of these cases include Proceedings on the Internet.



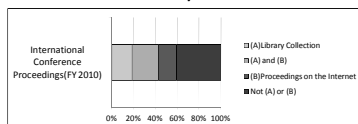
\*Published in two or more forms was classified into "Published in many forms" and duplication is not counted.

4

## 2. Publication forms of Proceedings and the library collection

### 2.2 Collection of Proceedings at the library

- We surveyed the library collection\* of the Proceedings written by JAEA researchers in FY2010.
- The JAEA library is collecting conference Proceedings in the fields of nuclear science and technology. However, JAEA library and other libraries hold only about 42% of the Proceedings. This demonstrates the difficulty of providing Proceedings only in a library collection.
- Providing the link to Proceedings on the Internet as well as the library collection are both important to increase access to the Proceedings.**



\*Library Collection:  
 ·JAEA Library  
 ·National Diet Library(NDL)  
 ·Cinii Books : Search information about books and journals held by university libraries in Japan

5

## 3. Management & providing the hyperlink

### 3.1 Providing the hyperlink to Proceedings

- At JAEA, researchers are requested to submit application forms before they make presentations or submit articles.
- Through this procedure, library staffs manage the information and make an authority file of author names, conference names, and journal/book names.
- The library is the gateway to the information on JAEA researchers' presentations/publications.
- We disseminate this information via the Internet using JOPSS (JAEA Originated Papers Searching System).
- In addition, we have started to provide hyperlinks to the full text of articles submitted by JAEA researchers with JOPSS (from July 2011 ).

6

### 3. Management & providing the hyperlink 3.1 Providing the hyperlink to Proceedings

- The bibliographic data of JAEA R&D results are released at JOPSS after checking them. The registration of DOI/URLs is performed every month.
- As few Proceedings papers have a DOI, URL have to be included. In addition, URL have to be added to the transactions in which the abstracts are published.
- As it stands, we provide 14,459 papers from conference presentations. Among them, 3,137 are hyperlinked to the full text or abstract (as of Oct. 2012).
- Providing the link by JOPSS allows direct access to the Proceedings on the Internet.

7

---

---

---

---

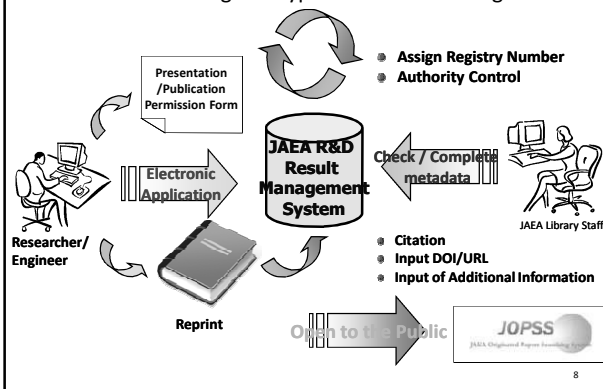
---

---

---

---

### 3. Management & providing the hyperlink 3.1 Providing the hyperlink to Proceedings



8

---

---

---

---

---

---

---

---

### 3. Management & providing the hyperlink 3.2 Surveys of Proceedings on the Internet

- A link to the Proceedings is not permanent because conference secretariats are often a temporary body and the URL is frequently changed.
- We surveyed the accessibility of the Proceedings on the Internet / Conference Website where JAEA researchers submitted papers in FY2010.
- 40% of the Proceedings are published on the Internet. Most are published on the conference website or a link to the Proceedings is provided.
- On the other hand, conference websites are likely to change URLs, and some of them are closed after the conference. 30% of the conference websites in FY2010 were not available.
- In addition, some of the Proceedings on the Internet are not open to the public. They permit access only to conference participants. These kinds of Proceedings are few, but access to full text is very difficult.

9

---

---

---

---

---

---

---

---

### 3. Management & providing the hyperlink 3.3 New approach to provide Proceedings information

- The link to Proceedings on the Internet as well as in the library collection is useful to improve accessibility to the Proceedings.
- We developed a prototype webpage that provides conference information, bibliographic information of Proceedings, and Proceedings on the Internet using Google Calendar.
- Many conference websites are deleted within a few years after the conference; therefore, an archive website would be useful. We use the Internet Archive (Wayback Machine (<http://archive.org>) to provide the conference information from the deleted conference websites.
- We use hyperlink to the JAEA Library/NDL/Cinii Books search result.

10

---

---

---

---

---

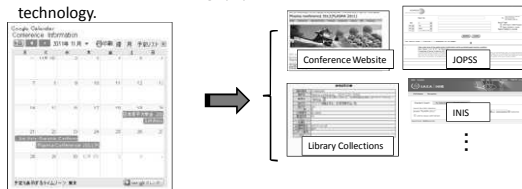
---

---

---

### 3. Management & providing the hyperlink 3.3 New approach to provide Proceedings information

- Some of the Proceedings papers are included in an institutional repository by the authors. An institutional repository can include Proceedings not available anywhere else. So, we provide the hyperlink to the search results of JAIRO (Japanese Institutional Repositories Online).
- In addition, we provide the hyperlink to the search results of INIS (International Nuclear Information System), which includes information from Conference Proceedings papers in the fields of nuclear science and technology.



11

---

---

---

---

---

---

---

---

### 4. Summary

- There are multiple publication forms of Proceedings, and the number of Proceedings available on the Internet are increasing. Using a link to Proceedings on the Internet as well as library collections can improve access to the Proceedings.
- The URLs of Proceedings on the Internet, especially those on conference websites, are more likely to change. When we provide the links to Proceedings, we need to periodically check the links.
- Some of the Proceedings on the Internet are not open to the public. An institutional repository will improve access to full text that is not available anywhere else.
- Using the Internet archive and institutional repository will probably improve the access to the full text of papers or paper information, and we would like to offer this information collectively.
- We would like to add an institutional repository in other country, such as Open DOAR.

12

---

---

---

---

---

---

---

---

## Envia: Environmental Information Discovery and Access

**Johanna Kieniewicz, Elizabeth Newbold, and Alan Sudlow**

Science Technology and Medicine, The British Library, United Kingdom

Envia is an online service being developed to provide co-ordinated access to environmental information - wherever it is and wherever you are. Working in partnership with Living With Environmental Change (LWEC), The British Library has developed Envia, a proof-of-concept service to demonstrate federated discovery and access of environmental information. Based upon extensive research into information use and needs across the environmental sciences sector, Envia will enable discovery and access of government reports, journal articles, PhD theses, datasets and other information held both within a British Library repository, as well as other distributed locations on the web. Although the service will have a British Library web presence, the aim is for Envia to be primarily accessed from wherever users of Environmental Science research are likely to be already. Over the next four years, what is now a proof-of-concept demonstrator will be transitioned to a sustainable service. For the Pilot phase, Envia has a focus on information relevant to those working on flooding; however, it is expected that this scope will expand incrementally over the coming years.

### Bionotes

**Johanna Kieniewicz**, STM Research and Engagement Manager, Environmental Sciences - Following a PhD in Earth and Planetary Sciences at Washington University in St Louis, she then went on to a position as an assistant professor at Denison University in the USA before moving to the UK and joining the Science Team at the British Library. She also has a background in art and has an abiding interest in the interface of art and science.

**Elizabeth Newbold** - After completing a degree in applied sciences and an MSc in Information Studies Elizabeth started a career as an information professional specialising in scientific, medical and technical information provision. Elizabeth has worked in a number of organisations and specialist information units, providing information research services to researchers in both the public and private sector before moving to the British Library in 2003 to manage the science, technology and medicine (STM) collections. Currently leading the STM Content & Collections team for the British Library her work focuses on collection and content development strategies and policies. Her interest in grey literature has developed over the years from her experiences working in organisations that were both users and producers of grey literature. Email: [Elizabeth.Newbold@bl.uk](mailto:Elizabeth.Newbold@bl.uk)

## Grey Literature, E-Repositories, and Evaluation of Academic and Research Institutes: The case study of BPI e-repository

**Maria V. Kitsiou and Vasileios Souvlidis**  
Benaki Phytopathological Institute, Greece

E-repositories are internet databases, in which the whole intellectual property produced by an Educational Foundation or a Research Institute can be gathered, classified, reserved and, of course, disseminated.

It is known that e-repositories are based on Open Access and Knowledge Dissemination concept, providing access-without restriction- to scientific information. But, it is not known that e-repositories can be proved useful and usable “tools” for the evaluation of Universities and Research Institutes.

In this study, we present the case of the BPI e-repository. The Library of the Benaki Phytopathological Institute (BPI) realizing the importance of the evaluation – necessary for funded projects- developed the BPI e-repository.

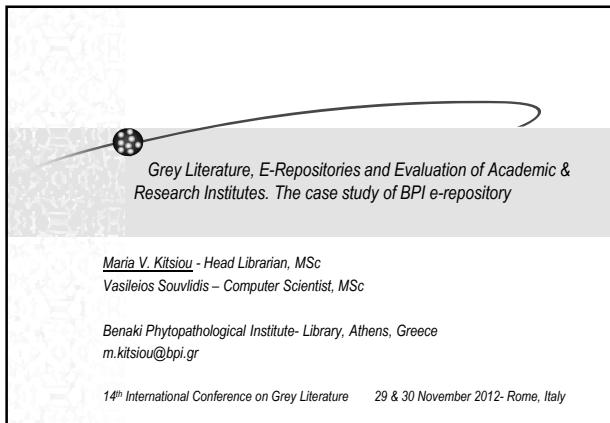
BPI e-repository has been planned and structured so in order to achieve the aims of the evaluation simultaneously in an “institution level”, in “scientific department level” (common intellectual responsibility) and in “researcher level” (individual intellectual responsibility). The scientific material contained in it has been *categorized* so in order to responds to the organizational structure and function of BPI and, also, to highlights in the best way the research activities of the BPI scientific community members. Its implementation based on *DSpace 1.7.0*, open source software suitable for digital archives management, that uses *OAI-PMH* (Open Archives Initiative Protocol for Metadata Harvesting) and Dublin Core Standard, that is suitable for the documents description. The flexibility of DSpace software allowed us to implement changes into the repository system, like fields addition. So, these changes made the description of all documents in different types (books or book chapters, papers, dissertations, BPI publications, patents, conferences, technical reports etc) more efficient and complete.

*Navigation* and *search* function in an user-friendly interface using five diverse ways, i.e. *type of material*, *title*, *author*, *year* and *subject*. We note that we implement a simple bilingual subject standardization system. The keywords used by the authors themselves in their papers or conferences proceedings make the subject standardization procedure and indexing difficult. Nevertheless, we are aiming at the implementation of an optimized bilingual subject index adopting and implementing *standards such as NLGSH or/and LC*..

The appropriate use and exploitation of the search results via e-repository – according to the experience having been acquired by this e-repository development process and function- can lead certainly to faithful conclusions concerning the evaluation of an Educational Foundation or a Research Institute.

### Bionote

**Maria V. Kitsiou** is born in Ioannina, Epirus, in north-western Greece. She holds a BSc in Archive and Library Science and a MSc in Library Management from Ionian University (Corfu, Greece). She has worked as a Librarian in a range of Libraries (Public Library of Lefkas, Special Library of Technical Camper of Greece, Academic Library of Technological Educational Institute of Ionian Islands). Since 2009 she has been working in the Library of Benaki Phytopathological Institute acting as a Head Librarian. Her interests include library assessment, quality systems implementation, information literacy- user education, grey literature management systems, e-repositories, open access, etc. Also, she has a publication in *Library Management* (vol. 29, issue 6/7, 2008) titled “Issues and perceptions for ISO 9000 implementation in Greek Academic Libraries”. Email: [m.kitsiou@bpi.gr](mailto:m.kitsiou@bpi.gr)



**Grey Literature, E-Repositories and Evaluation of Academic & Research Institutes. The case study of BPI e-repository**

Maria V. Kitsiou - Head Librarian, MSc  
 Vasileios Souvildis - Computer Scientist, MSc

Benaki Phytopathological Institute- Library, Athens, Greece  
 m.kitsiou@bpi.gr

14<sup>th</sup> International Conference on Grey Literature 29 & 30 November 2012- Rome, Italy

---

---

---

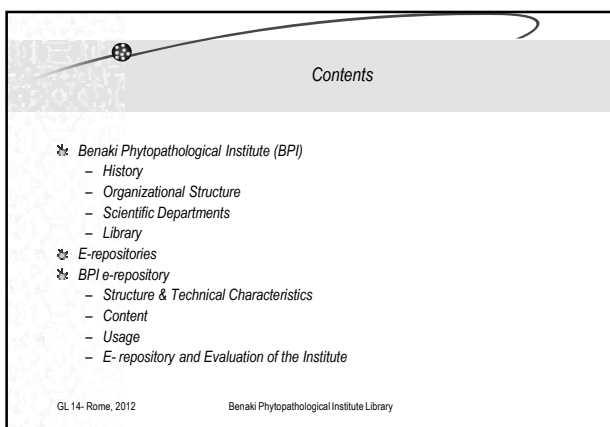
---

---

---

---

---



**Contents**

- ✧ Benaki Phytopathological Institute (BPI)
  - History
  - Organizational Structure
  - Scientific Departments
  - Library
- ✧ E-repositories
- ✧ BPI e-repository
  - Structure & Technical Characteristics
  - Content
  - Usage
  - E- repository and Evaluation of the Institute

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

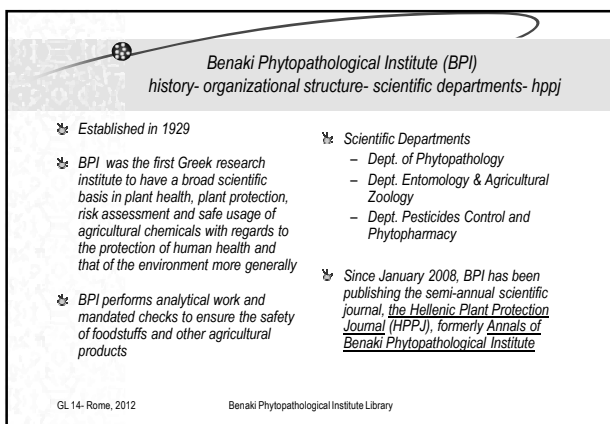
---

---

---

---

---



**Benaki Phytopathological Institute (BPI)**  
 history- organizational structure- scientific departments- hppj

- ✧ Established in 1929
- ✧ BPI was the first Greek research institute to have a broad scientific basis in plant health, plant protection, risk assessment and safe usage of agricultural chemicals with regards to the protection of human health and that of the environment more generally
- ✧ BPI performs analytical work and mandated checks to ensure the safety of foodstuffs and other agricultural products
- ✧ Scientific Departments
  - Dept. of Phytopathology
  - Dept. Entomology & Agricultural Zoology
  - Dept. Pesticides Control and Phytopharmacy
- ✧ Since January 2008, BPI has been publishing the semi-annual scientific journal, the Hellenic Plant Protection Journal (HPPJ), formerly Annals of Benaki Phytopathological Institute

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---



**Benaki Phytopathological Institute (BPI) Library**

- ✱ The Library of BPI has been in operation since 1930
- ✱ It is a specialized scientific library, dedicated to the subject of agriculture, intended to support and enhance the research activities of the BPI community
- ✱ In particular, it covers the subjects of phytopathology, plant protection, plant health, entomology, phytopharmacy and weed science
- ✱ Its collection consists of books, journals, grey literature, maps, digital and electronic media and databases
- ✱ The material is arranged according to national and international standards such as AACR2, DDC, UDC, NLGSH (National Library of Greece Subject Headings) and LCSH (Library of Congress Subject Headings)

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

**Benaki Phytopathological Institute (BPI) Library**

- ✱ Procedures concerning cataloguing as well as loan services are automated, using Abekt 5.6, an integrated library system that has been developed by National Documentation Center and installed in more than 2,100 libraries in Greece
- ✱ Quality system implementation
- ✱ Provided Services
  - Loans
  - Articles e-delivery
  - User education
  - E-index of journals & separated articles
  - **E-repository**

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

**E-repositories**

- ✱ E-repositories are internet databases, in which the intellectual property produced by an Educational Foundation or a Research Institute can be
  - Gathered
  - Classified
  - Reserved
  - Disseminated
- ✱ It is known that e-repositories are based on Open Access and Knowledge Dissemination concept, providing access to scientific information  
...but...
- ✱ it is not known that e-repositories can be proved useful and usable "tools" for the **evaluation** of Universities and Research Institutes

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

**BPI e-repository**  
structure, material contained, technical characteristics 1/2

- ✳ BPI e-repository has been planned and structured so in order to achieve the aims of the evaluation simultaneously,
  - in an "institution level"
  - in "scientific department level" (common intellectual responsibility)
  - in "researcher level" (individual intellectual responsibility)
- ✳ The scientific material contained in it has been categorized so in order to respond to the organizational structure and function of BPI and, also, to highlight in the best way the research activities of the BPI scientific community members

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

**BPI e-repository**  
structure, material contained, technical characteristics 2/2

- ✳ BOOKS
- ✳ BPI ANNUAL REPORTS
- ✳ BPI EDITIONS
- ✳ CONFERENCES
- ✳ FUNDED PROJECTS
- ✳ MEETINGS/ SEMINARS
- ✳ MONOGRAPHS
- ✳ MSc, PhD, Post Doc
- ✳ PATENTS
- ✳ PUBLICATIONS
- ✳ TECHNICAL REPORTS
- ✳ Characteristics
  - DSpace 1.7.0
  - OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)
  - Dublin Core Standard, that is suitable for the documents description
- ✳ Modifications
  - Fields addition

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

**BPI e-repository**  
fields/ records

a record consists of the fields:

- author(s)
- title
- date of issue
- publisher
- citation
- identifiers (ISSN, ISBN, DOI)
- type
- language
- subjects
- abstract
- description

... in numbers...

- ✳ 11 collections (1 more soon)
- ✳ 32 sub- collections
- ✳ 516 records
- ✳ 650 authors
- ✳ 1500 subjects
- ✳ a few months
- ✳ 1 librarian
- ✳ 1 computer scientist

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

**BPI e-repository**  
navigation, research, subject index

- Navigation and search function in an user-friendly interface using five diverse ways, i.e. type of material, title, author, year and subject
- The keywords used by the authors themselves in their papers or conferences proceedings make the subject standardization procedure and indexing difficult
- We note that we implement a simple bilingual subject standardization system
- We are aiming at the implementation of an optimized bilingual subject index adopting and implementing standards such as NLGSH or/and LCSH
- <http://83.235.16.144:8080/ispuien/>

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

**Conclusion**

*The appropriate use and exploitation of the search results via e-repository – according to the experience having been acquired by this e-repository development process and function- can lead certainly to faithful conclusions concerning the evaluation of an Educational Foundation or a Research Institute*

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

**Thank you for your attention!**

GL 14- Rome, 2012 Benaki Phytopathological Institute Library

---

---

---

---

---

---

---

---

"The GreyForum is a series of onsite and online courses, seminars, and workshops where grey literature provides common ground for information professionals in the process of knowledge transfer"

## Grey Literature and Information Ethics: Key Shared Concerns



Artwork by Jeroen Hoogstraten

### GreyForum 1.1

#### Workshop



#### Grey Literature and Information Ethics: Key Shared Concerns

#### Date

Thursday, 31 January 2013

#### Time

10:00 – 17:00 hours

#### Place

St Hilda's College  
Cowley Place, Oxford OX4 1DY  
[www.st-hildas.ox.ac.uk](http://www.st-hildas.ox.ac.uk)

#### Workshop Leaders

Jonathan Gordon-Till and  
Sylvia Simmons, InfoEthics UK  
Dominic Farace, GreyNet

#### Registration

£155 / €195 incl. Lunch, Taxes,  
and Workbook ISSN 2213-5735  
\*Special rates available online

#### Contact

TextRelease/GreyNet  
Grey Literature Network Service  
Javastraat 194-HS  
1095 CP Amsterdam, Netherlands

[info@greynet.org](mailto:info@greynet.org)  
[greynet.org/greyforumseries.html](http://greynet.org/greyforumseries.html)  
Phone/Fax +31 (0) 20 331 2420

*The first in this series is a workshop organized by InfoEthics UK in collaboration with GreyNet International and will be held in Oxford*

### Grey Literature and Information Ethics: Key Shared Concerns Driving organizational improvement in the digital information economy

Librarians and information professionals live and work in challenging economic times. Their organizations must be nimble, proactive and innovative as they emerge from the downturn. The knowledge worker's role is pivotal to ensure survival. Grey literature – viewed in the past by some as peripheral – is now ever more relevant. Now more than ever is the time to understand it.

GreyNet International and InfoEthics UK share key concerns facing the digital information community. A comparative analysis identifies: peer review • digital preservation • privacy • standardization • validation and • equitable pricing. Addressing these major issues with our Action Plan, you will be equipped to deal with emerging challenges in the digital information economy.

Jonathan Gordon-Till and Sylvia Simmons (InfoEthics UK) were founder members in 1992 of the Institute of Information Scientists Ethics Task Force, developing a pioneering Code of Professional Practice for information professionals. Dominic Farace founded the Grey Literature Network Service in 1992, which is dedicated to Research, Publication, Open Access, and Education in the field of Grey Literature.

**Grey Literature and Information Ethics** is aimed at information professionals, librarians, archivists, records and knowledge managers at all levels and in all sectors involved in the management of information resources.

**Outcomes** • *Effective identification and resolution of ethical dilemmas* • *Better decision-making and advisory skills* • *Better understanding of the role of professional Codes of Practice* • *Understanding risk management and compliance* • *Enhanced reputation for you and your organization* • *Improved due diligence and professional practice* •

#### Program Outline

09:30 Registration Desk Open	13:30 Case Studies
10:00 Introduction	14:45 Afternoon Tea
10:15 Ethical Decision Making	15:00 Drafting Consensus
11:30 Grey Literature Concerns	16:00 Review and Action Plan
12:45 Lunch	17:00 Close

\* Online Registration, <http://www.greynet.org/greyforumseries.html>

*GreyNet*

## Innovation, language, and the web

**Claudia Marzi**

Institute for Computational Linguistics, “Antonio Zampolli”,  
CNR, National Research Council, Italy

Language and innovation are inseparable. Language conveys ideas which are essential in corporate innovation; innovation would be nearly impossible if we did not have language. Language establishes the most immediate connections with our conceptualisation of the outside world, and it provides the building blocks for communication. The structure of language itself reflects its functional and communicative use. Communication takes place when there is a real information exchange process. Every linguistic choice is necessarily meaningful, and absolute variables involve the parallel construction of form and meaning.

From this perspective, language is not only structure, but a dynamic knowledge construction process as well. Knowledge transfer and innovation transfer are ubiquitous processes: knowledge extraction requires heterogeneous tasks related to the acquisition, from unstructured textual data in digital format, of structured and classified information relating to research topics.

In the full version of this approach, emphasis will be laid on the mechanisms underlying language processing and communicative interaction, outlining knowledge retention and retrieval processes.

The spread of Internet has enabled development of better bibliographic scientific databases with significantly improved capacity for storage and retrieval. In recent years, web searching has become the default mode of highly innovative information retrieval, though the main sources of digital information are unstructured or semi-structured documents. Information relating to developments in scientific research is collected in the form of abstracts or full publications, in large and growing bibliographic repositories. Considering the web as a corpus makes it possible to investigate how words are used to describe innovation, and how innovation topics can influence word usage and collocational behaviour. Investigation of corpora is concerned with the description of use and structure of language, by inquiring linguistic phenomena such as, co-occurrence distributions, collocational variability, derivational productivity, neologism coinage.

This will bring into focus the dynamic interplay between lexical creativity and innovative pragmatic contexts, thus blurring the traditional dichotomy between knowledge of language and its use.

In particular, the work will focus on how words and language structures become vehicle for knowledge generation and innovation transfer, and how research data, research results and widely-distributed dissemination papers can support and enhance future research.

### Bionote

---

**Claudia Marzi** is a research fellow at the Institute for Computational Linguistics “A. Zampolli” (ILC), National Research Council (CNR), Pisa. PhD in Acquisitional and Computational Linguistics, University of Pavia, in progress. Laurea degree in Modern Languages (English) at Pisa University in 1998, with the dissertation “The power of words: language creativity in Edgar Allan Poe's narrative”. Programme Coordinator of the European Science Foundation Research Networking Programme “The European Network on Word Structure. Cross-disciplinary approaches to understanding word structure in the languages of Europe”. Member of board at Institute for Computational Linguistics. Main areas of interest: Computer modelling of the Mental Lexicon; Second language acquisition; Child language; Document and knowledge management. Email: [claudia.marzi@ilc.cnr.it](mailto:claudia.marzi@ilc.cnr.it)



**INNOVATION, LANGUAGE, and the WEB**

**Claudia Marzi**

**Institute for Computational Linguistics, "Antonio Zampolli" –  
Italian National Research Council**

**University of Pavia – Dept. of Theoretical and Applied Linguistics**





---

---

---


---

---


---

---

---



**INTRODUCTION**




Language conveys ideas which are essential in corporate innovation; innovation would be nearly impossible if we did not have language.

Language makes us shape and refer to things, events, and concepts.

Communication takes place when there is a real information exchange process:

- every linguistic choice is necessarily meaningful
- language is in this sense a dynamic communicative and interactive process




---

---

---


---

---

---

---



---



**RESEARCH GOAL**

The communicative ability connects people into information-sharing network and information allows people to expand their knowledge. In this context, our goal is to:

- ➔ focus on how words and language structures become vehicle for knowledge generation, and in particular for innovation transfer.
- ➔ focus on lexical creativity, and its dynamic interplay with innovative contexts


---

---

---


---

---

---

---


---



## WEB AS A CORPUS


**The WEB:**

- has become a primary meeting place for **information exchange**;
- has evolved into a primary **resource** for lexicographers and linguists;
- Is a source of machine readable texts for corpus linguists and researchers in the fields of Natural Language Processing (NLP), Information Retrieval and Text Mining



➔

**WEB searching** as default mode for information retrieval, though the main sources of digital information are unstructured or semi-structured text materials



---

---

---


---

---

---

---

---




## CORPUS LINGUISTICS

The meanings of words can change over time/discourse, and words can take on new senses

➔ **Large corpora** – large and growing repositories - to investigate:

- how words are used to describe innovation
- how innovation-driven topics can influence word usage and collocational behaviour

➔ **Corpus-based studies** → lay emphasis on how word usage and word formation processes conveys differences in **context, domain** and **purpose**



---

---

---


---

---

---


---

---



## RESEARCH RATIONALE

- ★ The meanings of words can change over time, and words can take on new senses
- ★ Words with emergent novel senses reflect an extension of use from one domain to another
- ★ Semantically ambiguous words (polysemy) can be disambiguated by defining the context – what people do in spoken language (pragmatic competence)
- ★ Analysis on language productivity offer the opportunity to investigate how words become vehicle for innovative knowledge generation and transfer



---

---

---


---

---

---



---

---



## METHODOLOGY

---

-  Genre-oriented and stylistically heterogeneous texts are taken:
- Bibliographic (domain-specific) database
- Type coherent multidisciplinary database
  
- Search web engine and large corpus query tools are used:
- Generic search engine
- Lexicography query system
  
-  Domain-specific vs. generic-purpose texts offer materials for terminology exploration

---


---

---

---

---


---



EC


## EXPERIMENTAL EVIDENCE

---




By comparing reference corpora – both specific and generic – with the web, examples of polysemy are given:

- ✓ **Imaging** = brain imaging techniques vs. visual representation → **ambiguous**
- ✓ **Neuro** = referred to diagnosis, medicine and surgery. → **coherent**
- **Neuroimaging** → **coherent**
- ✓ **Storage** = memory capacity vs. containing units → **ambiguous**
- ✓ **Memory** = process of information encoding and retrieving → **coherent**



What about a polysemous term like **retention**? What domain does it belong to? What is the meaning and what concepts does it convey?



EC

## EXERCISES

---

---

---

---

---

---

## DISCUSSION and CONCLUSION

- ★ Lexical co-occurrences and collocations can be of considerable help in retrieving text materials which are relevant to a specific domain of interest; however they are often helpless in telling genres purposes or intended readerships
- ★ Bibliographic domain-specific databases are developed to offer materials to a specialised readership, thus providing highly-selected, well-targeted documents
- ★ We investigate the hypothesis that specific word-formation processes and neologisms correlate significantly with text genres and intended readership, thus providing a convenient way to track down well-targeted, highly technical repositories of openly available text materials

---

---

---

---

---

---



## Open Grey for Natural Language Processing: A ride on the network

**Gabriella Pardelli, Sara Goggi, Manuela Sassi**

Istituto di Linguistica Computazionale, “Antonio Zampolli”, ILC-CNR, Italy

Open Access is the key in the development of Information Society (IS), a new method for sharing scientific resources which influences the dynamics of creation and dissemination of knowledge. In order to share and spread this knowledge ever more sophisticated digital devices are tuned up while scientific institutions and associations are lately committed to the creation of dedicated repositories with the intent of giving wide visibility to their resources.

There is more: sharing open access information does not only mean retrieving objects of digital nature from the origins but also digitally reproducing source material from the far and recent past. The aim of this paper is to introduce the Open Access movement for Natural Language Processing (NLP) by means of a wide range of open access Grey Literature documentation present on the web. In 2008 Robert Dale, in the last issue of volume 35 of Computational Linguistics said: “There are a number of definitions of the term ‘open access’ in circulation, but almost all share the key principle that scientific literature should be freely available for all to read, download, copy, distribute, and use (with appropriate attribution) without restriction”. At first glance it might seem that the Open Access movement has gradually become more influential in the field of language technology by building repositories accessible through the network. Today's digital archives are niches of intellectual production spread by means of a wide range of documents (such as journal articles and proceedings) which, paradoxically, the search engines do not always reach. The use of inappropriate terms in the formulation of queries and the fragmentation of repositories in this area of investigation does not allow to retrieve information on a large scale.

The definition of open access is rather tricky, as Merkel-Sobotta says in 2005: “ ‘Open access’ means many different things to many different people. To use an example from US politics: it is as difficult to be anti-choice as it is to be anti-life. In the flux of ideas generated by the new and rapidly developing phenomenon of web publishing, open access proponents were able to convince others that traditional publishers were “anti-open access” or even anti-access, period. Tested against the realities of e-publishing, this did not last very long”.

The following are a few examples of significant Open Access repositories of our field: the Association for Computational Linguistics (ACL) built a rich repository called ACL Anthology, a digital archive of research papers in Computational Linguistics.

This archive traces down the history of Computational Linguistics from first research of the '60s by retrieving and putting on the web the articles published in the proceedings of the most important international conferences of the field (i.e. COLING series).

Since some time several conferences publish their contributions as open access documents: the Global Wordnet Conference and the Language Resources and Evaluation Conference (LREC), just to make a couple of examples.

Machine Translation Archive is an electronic repository and bibliography of articles, books and papers on several topics in the field of machine translation, computer translation systems and computer-based translation tools. This archive contains knowledge: its documents, accessible by everyone, provide an historical overview of automatic translation which might turn out to be very useful both for experts and non-experts of the field.

Both conferences and workshops articles are available on line; the material is widely spread over the web. As a matter of fact, the available search engines – though very sophisticated – are not nowadays able to provide a comprehensive plan of open access resources for Language Technology.

The full paper, after a first introductory section, will be organized in two sections:

1) the first dedicated to the methodology for searching and tracing open access resources and to the criteria for analyzing and selecting the online documentation;

2) the second devoted to a description of the state-of-the-art of Open Access Grey Literature material in a statistical and thematic scenario.

As things stand, standardization of computational systems interconnected by links and tools of various nature allowing Internet users to easily retrieve the information that the web naturally makes available would then be essential.

## References

Dale R., LastWords What's the Future for Computational Linguistics? Computational Linguistics, 2008, Volume 34, Number 4, Pages 621-624.

Merkel-Sobotta E., Elsevier and Open Access, Neuroinformatics, 2005, Volume 3, Pages 5-9.

<http://www.springerlink.com/content/x66772m265840111/fulltext.pdf>

Ranger Sara L., Grey Literature in Special Library: Access and the Use, Publishing Research Quarterly, 2005, Volume 21, Number 1, Pages 53-63.

<http://www.aclweb.org/anthology/J/J08/J08-4008.pdf>

<http://www.ling.ohio-state.edu/acl08/cfp.html>

<http://www.mt-archive.info/>

<http://www.aclweb.org/anthology/>

<http://www.cfilt.iitb.ac.in/gwc2010/>

<http://www.lrec-conf.org/proceedings/lrec2010/index.html>

<http://puma.isti.cnr.it//index.php?langver=it>

<http://www.greynet.org/greytextarchive.html>

## Bionotes

**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](mailto:gabriella.pardelli@ilc.cnr.it)

**Manuela Sassi** graduated in Foreign Languages and Literature at Pisa University, 110/110 cum laude. Since 1974 she has been working in Pisa at the Institute for Computational Linguistics of the National Research Council. Her interests and experiences range from linguistic to textual data processing and in providing linguistic resources on-line. She has been responsible for many national projects and has participated in numerous international projects.

Email: [manuela.sassi@ilc.cnr.it](mailto:manuela.sassi@ilc.cnr.it)

Open Grey for Language Technology: a ride on the network

*Gabriella Pardelli, Manuela Sassi, Sara Goggi*  
*Istituto di Linguistica Computazionale "Antonio Zampolli", ILC*  
*Consiglio Nazionale delle Ricerche CNR- Pisa, Italy*

**GL14 Fourteenth International Conference on Grey Literature**  
**National Research Council, Rome, Italy 29-30 November 2012**

---

---

---

---

---

---

---

---

Starting Questions

- ❑ Where to search for Language Technology (LT) documents?
- ❑ How to retrieve LT documents?
- ❑ Is the extensive experience in LT important for retrieving information?

GL14 Fourteenth International Conference on Grey Literature  
National Research Council, Rome, Italy 29-30 November 2012

---

---

---

---

---

---

---

---

The first decade of the new millenium has suddenly witnessed first the growth and then the rapid increase of the so-called "social networks" which totally transformed the way information is transmitted: nowadays the World Wide Web looks like an enourmous collection of documents inter-connected and linked to the various search engines by sharing the same paradigm (Web 2.0).

The selection, conservation and storage of digital content apparently makes the users' fruition easier : but is this assumption really true?

To formulate appropriate and effective queries for a search is a difficult task for users and requires a careful terminological selection for obtaining the most from an Information Retrieval system

↓

Information Retrieval is the academic discipline that studies the methodologies, tools, techniques, and languages for searching and retrieving relevant data for an information need.

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012

---

---

---

---

---

---

---

---

## Where to search for Open Access Language Technology documents?

When a user tries to retrieve information on a given topic from online repositories, there are several possibilities to formulate a query; for instance, given the query "Language Technology", the web replies with about 878.000.000 results in 0,26 seconds (September 25, 2012 at 4.30 p.m.):

### Academic articles for language technology

...of the state of the art in human language technology - Cole - Cited by 577  
 ...Stirring up trouble about language, technology and ...-Postman-Cited by 158  
 ...Information extraction as a core language technology - Wilks - Cited by 69

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012



1. [Language technology](#) - Wikipedia, the free encyclopedia on wikipedia.org/wiki/Language\_technology - [Traduci questa pagina](#)  
 Language technology is often called human language technology (HLT) or natural language processing (NLP) and consists of computational linguistics (or CL) ...
2. [Welcome to Language Technology World](#) - IT World, [www.it-world.org](#) - [Traduci questa pagina](#) 4 Feb 2012 - A portal on the range of technologies that deal with human language. News, conferences, projects, organisations, systems, and resources.
3. [CLIL - Language and Information Technology](#) [www.cclil.it](#) - [Traduci questa pagina](#) L'analisi del linguaggio è diventata un fattore di successo dei nostri clienti. Language Specialist multilingua di CLIL lavorano in più di 30 lingue diverse.
4. [Language Learning & Technology - Home](#) [llt.nyu.edu](#) - [Traduci questa pagina](#) Online journal devoted to technology and language education research for foreign and second language educators. Full text of articles available.
5. [UNITS Human Language Technology and Interfaces](#) [www.unite.it/interact/](#) [Human language technology and interfaces](#) Le Tecnologie del Linguaggio (Human Language Technologies, HLT) ci permettono oggi di interagire a voce con vari servizi automatici, ad esempio per ...
6. [\[PDF\] Language Technology A First Overview - DFKI](#) [www.dfki.de/~human/LT.pdf](#) - [Traduci questa pagina](#) Formato file: PDF/Adobe Acrobat - [Visualizzazione minima](#) di IT Uskorell - [Clicca da 2](#) - [Analisti correlati](#) Language technologies are information technologies that are specialised for dealing ... are also often subsumed under the term Human Language Technology.
7. [DFKI Language Technology Lab](#) [www.dfki.de/lt](#) - [Traduci questa pagina](#) Die Deutsche Forschungszentrum für Künstliche Intelligenz GmbH mit Sitz in Kaiserslautern und Saarbrücken ist auf dem Gebiet innovativer ...
8. [CMU - Language Technologies Institute](#) [www.lti.cs.cmu.edu](#) - [Traduci questa pagina](#) CMU/LLI offers MS and PhD programs in Language and Information Technologies.
9. [Language Technology](#) [www.lang-tech.org](#) - [Traduci questa pagina](#) LangTech is the european forum dedicated to communities and organisations involved in the development, deployment and exploitation of Language and ...
10. [Immagini relative a language technology](#) - [Sei sulla pagina](#) non appropriata.
11. [FBK Human Language Technology](#) [hlt.fbk.eu](#) - [Traduci questa pagina](#) FBK - Fondazione Bruno Kessler. Human Language Technology. FBK > IT. News. 10 Sep 2012. Demo Paper accepted at ISWC 2012. 03 Sep 2012 ...

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012

## My findings...

From this generic query it is possible to retrieve only:

- ✓ 9 portals
- ✓ 2 open access publications (Cole and Wilks)
- ✓ 1 review (Portman)

But only 2 documents from this set satisfy the user.....

Lets' try with queries formulated by an expert user:

- 1 Query: ACL Anthology



- 2 Query: Language Technology

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012

The web answers... ..

<http://image.google.com/download/TECHNOLOGY.pdf>  
Google custom search

Over 11,800 results (0.16 seconds)

**Language Technology and Society Richard Sprockel**  
Language, Technology, and Society. Richard Sprockel. (Oxford Health & Science University). Oxford: Oxford University Press, 2010. xiv+285 pp. hardbound. ...  
scholar.google.com/citations?view\_op=view\_citation&hl=en&user=sprockel

1. **Building a social network from human language technology**  
Forman the 10th Annual Meeting on Empirical Methods in Natural Language Processing (EMNLP) 2010, December 26-31, 2010, Cambridge, Massachusetts, USA. ...  
www.aclweb.org/anthology/EMNLP-2010/
2. **Computational Language Technology Applications (CLTA)**  
The Australian Language Technology Association (ALTA) was founded at the 18th Australasian Natural Language Processing Workshop in Canberra, ...  
aclweb.org/anthology/australian-language-technology-association
3. **Emerging Language Technology: A Practical Overview**  
Forman the 10th Annual Meeting on Empirical Methods in Natural Language Processing (EMNLP) 2010, December 26-31, 2010, Cambridge, Massachusetts, USA. ...  
www.aclweb.org/anthology/EMNLP-2010/
4. **CLTA: Chinese Language Technology Applications**  
Forman the 10th Annual Meeting on Empirical Methods in Natural Language Processing (EMNLP) 2010, December 26-31, 2010, Cambridge, Massachusetts, USA. ...  
aclweb.org/anthology/EMNLP-2010/
5. **Building Language Technology: Where Do We Go From Here?**  
Forman the 10th Annual Meeting on Empirical Methods in Natural Language Processing (EMNLP) 2010, December 26-31, 2010, Cambridge, Massachusetts, USA. ...  
aclweb.org/anthology/EMNLP-2010/
6. **Building Language Technology: Where Do We Go From Here?**  
Forman the 10th Annual Meeting on Empirical Methods in Natural Language Processing (EMNLP) 2010, December 26-31, 2010, Cambridge, Massachusetts, USA. ...  
aclweb.org/anthology/EMNLP-2010/

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012

---

---

---

---

---

---

---

---

---

---

...

10 open access Grey Literature documents found!

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012

---

---

---

---

---

---

---

---

---

---

**How to retrieve LT documents**

Given the fact that the enormous amount of data available on the web is difficult to query from a semantic point of view, the human interpretation is always needed - - but which are the assumptions/conditions for making an effective query?

- being updated on the state-of-the-art;
- being skilled at navigating on the web portals;
- being able to understand the data.

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012

---

---

---

---

---

---

---

---

---

---

### "Noise" or "Silence"...

Performance of an information retrieval system can be measured by the following coefficients:

- ✓ *Precision*: proportion of relevant data retrieved from the total data retrieved
- ✓ *Recall*: extent of relevant data retrieved from the total data relevant in the database.

These coefficients measure two different factors:

- *Noise* = non-relevant data retrieved
- *Silence* = relevant data that have not been retrieved from the data base.

Retrieval models compute the degree to which certain elements answer to a query: a good model should be able to maximize recall and precision and minimize, respectively, "silence" and "noise".

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012

---

---

---

---

---

---

---

---

### Concluding Remarks

The web is both a knowledge repository and a knowledge dispenser



- Create innovative paradigms for information retrieval
- Establish features for semantic search on web portals
- Achieve precision & recall

**Nowadays knowledge extraction is only possible if:**

1. the know-how of the state-of-the-art is updated
2. appropriate queries are formulated by selecting effective terms from the dedicated portals.

GL14 Fourteenth International Conference on Grey Literature National Research Council, Rome, Italy 29-30 November 2012

---

---

---

---

---

---

---

---

## IRPPS Editoria Elettronica: An electronic publishing web portal based on Open Journal Systems (OJS)

**Marianna Nobile and Fabrizio Pecoraro**

Institute for Research on Population and Social Policies, IRPPS, Italy

In the last decades the development of sustainable economic models based on open source technologies for the management and dissemination of publishing activities represents a great opportunity for research institutions to improve the diffusion of their scientific information. The introduction of “*electronic publishing*” is leading libraries to play a key role in the improvement of innovative systems for the dissemination of scientific research results, with the aim of increasing the quality of products and reduce cost of publications. This is particularly evident considering the widespread diffusion of open access journals published by Research Institutions, Universities, etc.

Since its foundation, in 1981, the Institute for Research on Population and Social Policies (IRPPS) has carried out its publishing activities with the aim of diffusing its scientific results. An important and significant part of its editorial production includes traditional grey documents as *Working papers* and *Monographs*, both systematically published, while other editorial products (conference proceedings, reprints) are published when scientific events are organized by the Institute. There are also documents of particular historical interest as *the First report on demographic situation in Italy* published for the first time by the Institute in 1985, *The Italian families in the eighties* (1991) or *The Atlas of an aging population: twenty years of evolution in the provinces and in Italian municipalities* (1995). In this perspective, the IRPPS is reorganizing its scientific editorial activity to manage its in-house publications. *Open Journal Systems (OJS)* [1] is a suitable software that manages both editorial activities and open access scientific publications. In particular it supports the editorial workflow, from manuscript submission to publication including the peer-review process, it facilitates the e-publishing of different types of editorial products, from serials to monographs, thus allowing the reduction of journal's operating costs. At the moment the “IRPPS Editoria Elettronica” portal [3] publishes two series: *IRPPS Monografie* and *IRPPS Working Papers*. Besides essays, *IRPPS Monografie* features the proceedings of conferences organized by the Institute along with digital versions of already published works. Moreover, the editorial working group has defined the editorial policies to be adopted for both series identifying the different roles to be assigned in the publishing process, such as the editor, section editor, reviewers, etc. Finally, to improve the quality and visibility of the IRPPS scientific research various indexing procedure requests has been carried out in more closely disciplinary archives (i.e. Google Scholar, RePEc, etc...).

### References

- [1] Open Journal Systems (OJS). Available from: <http://pkp.sfu.ca/ojs>
- [2] Willinsky, J. (2006). The access principle: The case for open access to research and scholarship. Cambridge, MA: MIT Press.
- [3] IRPPS Editoria Elettronica. Available from: [www.irpps.cnr.it/e-pub](http://www.irpps.cnr.it/e-pub)

### Bionotes

**Marianna Nobile** graduated in Linguistics, Faculty of Letters and Philosophy at the University of Rome La Sapienza. She had an internship in the library of the Senate of the Republic (Giovanni Spadolini Library), where she worked with the Books Acquisition. Currently she is collaborating with the library of the Institute for Research on Population and Social Science of the Italian National Research Council (CNR) and she is involved in the activities of indexing electronic resources as well as in the development of a e-publishing service. Email: [marianna.nobile@irpps.cnr.it](mailto:marianna.nobile@irpps.cnr.it)

**Fabrizio Pecoraro.** Degree in Computer Engineering in Rome and Philosophy Doctorate in Bioengineering at the University of Bologna. During his doctorate studentship period he also held the position of assistant researcher at the University of Strathclyde, Glasgow, Scotland. Since 2007 he works as a researcher at the National Research Council – Institute of Research on Population and Social Studies, Rome Italy – where his research activities mostly focus on the following aspects: business process analysis, development of conceptual models based on standard of clinical data such as HL7 and CDISC, design and development of information systems and definition of relational databases. Email: [f.pecoraro@irpps.cnr.it](mailto:f.pecoraro@irpps.cnr.it)

## **IRPPS Editoria Elettronica:** *an electronic publishing web portal based on Open Journal Systems (OJS)*

Marianna Nobile and Fabrizio Pecoraro

National Research Council  
Institute for Research on Population and Social Policies, Rome – Italy

### Aim

Outline the e-publishing developed by the IRPPS service (called Editoria Elettronica in Italian) that intends to:

- reorganize the institute scientific editorial activity
- manage its in-house publications
- diffuse its scientific results



Fourteenth International Conference on Grey Literature - Tracking Innovation through Grey Literature  
29-30 November 2012, Rome, Italy



### Content

In particular, it poster describes:

- IRPPS editorial products
- The platform used to develop the service (Open Journal Systems)
- The system publishing process and the web portal developed



Fourteenth International Conference on Grey Literature - Tracking Innovation through Grey Literature  
29-30 November 2012, Rome, Italy





# FIND THE PIECE THAT FITS YOUR PUZZLE



## THE GREY LITERATURE REPORT FROM THE NEW YORK ACADEMY OF MEDICINE

Focused on health services research and selected public health topics, the Report delivers content from over 750 non-commercial publishers on a bi-monthly basis.

Report resources are selected and indexed by information professionals, and are searchable through the Academy Library's online catalog.

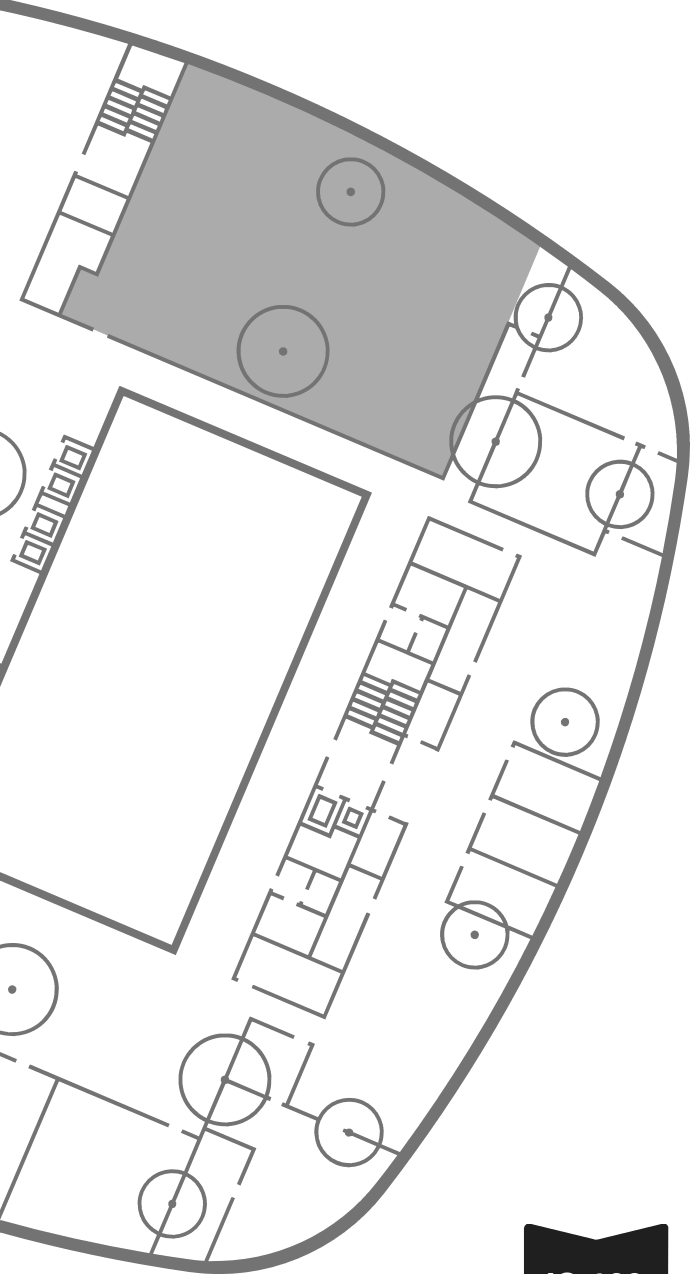
Let us help you put it all together; subscribe to the Grey Literature Report today!

For more information visit our website: [www.greyliterature.org](http://www.greyliterature.org)  
or contact us at: [greylithelp@nyam.org](mailto:greylithelp@nyam.org)

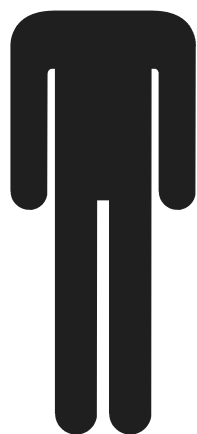


**The New York  
Academy of Medicine**

*At the heart of urban health since 1847*



IQ 166



## National Technical Library

**National Technical Library** (hereafter referred to as “NTK”) is a central professional library open to public, which offers a unique collection of 250 thousand publications freely accessible in open circulation. Its holdings form the largest collection of Czech and foreign documents from technology and applied natural sciences as well as associated social sciences. It contains a total of 1,2 Mil. volumes of books, journals and newspapers, theses, reports, standards, and trade literature in both printed and electronic forms. Besides its own collection, parts of Central Library of the CTU in Prague and Central Library of the ICT holdings are accessible in NTK.

For detailed information on the National Technical Library visit <http://www.techlib.cz/en/>

94 mm

As corresponds to its statutes, NTK manages – among others – the project of building the **National Repository of Grey Literature**.

The project aims at gathering metadata and possibly full texts of grey documents in the fields of education, science and research. The NTK supports an education in the field of grey literature through annual seminars in the Czech Republic.

For more information on the National Repository of Grey Literature visit our project Web site <http://nrgl.techlib.cz/> and for a search <http://www.nusl.cz/>

**NTK** Univers

Národní technická knihovna  
National Technical Library

NU  
SL

## Grey literature partnership network in the Czech Republic

**Petra Pejšová and Hana Vyčítalová**  
National Technical Library, Czech Republic

The poster presents grey literature partnership network in the Czech Republic. Partnership network used for SIGLE cooperation was dissolved at the beginning of 2005. New and broader partnership network was started to create at the end of 2009. In August 2012, there are 89 data producers from research and science, education, culture and also enterprise, namely research institutes, universities, libraries etc. Partnership network is built to support National Repository of Grey Literature (NRGL).

Grey literature data producers can choose from four technical possibilities how to collaborate with NRGL. All data are available as a central user interface at <http://www.nusl.cz> in both a Czech and English version. National Technical Library (NTK) as a promoter supports partnership network in many ways. NTK runs informative webpages, e-mail conference, publishes manuals and methodologies, performs presentations and trainings and every year organizes full day seminar with conference proceedings.

### *Bionote*

---

**Petra Pejšová** studied information science and librarianship at Charles University. She works as an information specialist in the State technical Library, Czech Republic. Actually she is leading a project Digital Library for Grey Literature – Functional model and pilot. Email: [petra.pejsova@techlib.cz](mailto:petra.pejsova@techlib.cz)

**NTK**

NTK 14.0037°N, 14°22'22.001"E  
Národní technická knihovna  
National Technical Library

210 mm

***Grey literature partnership network in the Czech Republic***

Fourteenth International Conference on Grey Literature

National Research Council, Rome, Italy 29-30 November 2012

Petra Pejšová  
National Technical Library, Czech Republic




---

---

---

---

---

---

---

---

**NTK**

NTK 14.0037°N, 14°22'22.001"E  
Národní technická knihovna  
National Technical Library

210 mm

**Starting points**

- **Reason:** data for the National repository of grey literature (NRGL)
- **Managed:** by the National technical library
- **Start:** at the end of 2009
- **Goal:** support research, development and education

---

---

---

---

---

---

---

---


**NTK**

NTK 14.0037°N, 14°22'22.001"E  
Národní technická knihovna  
National Technical Library

210 mm

**Today state**

- **Now 91 institutions:**
  - Universities
  - Institutes of Academy of Sciences
  - Research organizations
  - Libraries and museums
  - Ministries




---

---

---

---

---

---

---


---

**NTK**  
Národní technická knihovna  
National Technical Library

## Records

---

- Search interface  
<http://www.nusl.cz>
- Over 203 thousand of records
- Information about availability
- Search for mobile devices



---

---

---

---

---

---

---

---

**NTK**  
Národní technická knihovna  
National Technical Library

## Modes of cooperation

---

1. Connection of local database to NRGL search interface
2. Harvesting of records from local database into NRGL digital repository (INVENIO)
3. Installation of INVENIO as a local database which is harvested into NRGL digital repository
4. Submitting records and fulltexts directly into NRGL digital repository (INVENIO)
2. – 4. variant are also indexed into NRGL search interface

---

---

---

---

---

---

---

---

**NTK**  
Národní technická knihovna  
National Technical Library

## Support

---

- Informative webpages with news and RSS  
[http://nrql.techlib.cz/index.php/About\\_NRGL](http://nrql.techlib.cz/index.php/About_NRGL)
- Every year workshops  
<http://nrql.techlib.cz/index.php/Workshop>
- Moderated electronic conference
- Standards, manuals and recommendations
- Guarantee of long-term storage and access to data

---

---

---

---

---

---

---

---

**NTK**  
Národní technická knihovna  
National Technical Library

## Legal framework

- Legal expertise: Digital processing of the so-called grey literature for the NRGL
- Rules for legal publishing of digital documents in the NRGL
- Each work is protected by copyright law
- Selective choice of a license  
Creative Commons  
<http://www.creativecommons.cz>

**NTK**  
Národní technická knihovna  
National Technical Library

More information:

<http://nrql.techlib.cz/index.php/Cooperation>

Contact:

Petra Pejšová

[petra.pejsova@techlib.cz](mailto:petra.pejsova@techlib.cz)



## Building the evidence base to support food and feed safety risk assessments in Europe

**J. Richardson, L. Congiu, C. Morganti, P. Pirro, E. Aissa, S. Noorani, D. Lefebvre, and D. Verloo**  
EFSA, European Food Safety Authority, Italy

The European Food Safety Authority (EFSA) was established to assess risks associated with the food chain. EFSA's risk assessment work contributes to improving food safety in Europe and to building public confidence in the way risk is assessed. Risk assessment is a specialised field of applied science that involves reviewing scientific data and studies in order to evaluate risks associated with certain hazards. In order to ensure the risk assessment process is robust and transparent EFSA is using systematic literature review principles for the identification, selection and reporting of scientific studies used as the basis for risk assessments. The identification of scientific studies and risk assessments carried out in the Member States by competent authorities is critical for the risk assessment process. These reports, working papers and dissertations are often more difficult to identify and their identification is resource intensive and requires substantial topic expertise, since these types reports are less consistently archived and indexed in bibliographic databases. As a consequence EFSA in consultation with its Focal Point Network developed the Information Exchange Platform (IEP) to facilitate access to ongoing and recently completed risk assessments produced by Member State competent authorities. Building on experience gained from the IEP application; EFSA is now developing a repository for scientific documents to support food and feed safety risk assessments. The remit of EFSA is broad and encompasses toxicology, food and feed safety, nutrition, animal health and welfare, plant protection and plant health, environmental safety, assessment of regulated products and health claims, agrofood innovation, biological hazards, chemical contaminants, food consumption and emerging risks. The repository will be open to the scientific community and other European stakeholders to facilitate the upload and sharing risk assessments, research findings and reports covering these topic areas. The adoption of international metadata standards and faceted classification will ensure that risk assessors can quickly identify and access relevant reports. It is hoped that the EFSA repository will become a comprehensive evidence base for the European risk assessment community.

### Bionote

**Jane Richardson** is a database manager at the European Food Safety Authority (EFSA). She has a PhD in biological sciences and has experience in software development to support radiology services. More recently she has worked in the field of public health and the development of data management procedures and systems for epidemiology and exposure assessments. She is currently responsible for the implementation of the EFSA library web portal with a focus on applications and repositories to support systematic literature reviews and scientific risk assessments in food and feed safety. Email: [jane.richardson@efsa.europa.eu](mailto:jane.richardson@efsa.europa.eu)

## Tools and Resources Supporting the Cultural Tourism

**Eva Sassolini, Alessandra Cinini, Stefano Sbrulli, and Eugenio Picchi**

Istituto di Linguistica Computazionale, ILC, Italy

This paper presents the use of NLP techniques (text mining, text analysis) for development of specific tools that allow us to create linguistic resources related to the domain of "cultural heritage". In particular the project "The online dissemination of the historical artistic and scenic, regional heritage" was born in the framework of the collaboration between the Pisa ILC-CNR and the APT Basilicata (i.e. Agenzia di Promozione Territoriale della regione Basilicata) for experimenting and implementing strategies for the promotion and dissemination of regional heritage, particularly in initiatives related to tourism field.

The ILC contribution consists in the definition of a model of linguistic analysis of texts, with automatic extraction of semantic information and terminology to be used for the categorization of materials and their guided browsing.

We have experimented a particular methodology, dividing the automatic acquisition of texts and consequently of reference corpus, in two phases. The first phase aims at the extraction of documents from lists of links provided by experts worked with us to the project, for the creation of a repository of documents, whose selection criterion is geographical. In this case the documents can be considered as "grey literature" produced by the regional entities. On the basis of documents extracted from the web, by automatic spiders developed within TextPower technology, automatic parsers create the reference corpus for the cultural heritage domain. Relevant information and semantic concepts are then extracted from this corpus. All these semantically relevant elements (such as proper names, names of institutions, names of places, and other relevant terms) have been used as basis for further acquisitions of documents from heterogeneous sources, by using specialized crawlers that work on a bulk of text materials available on-line. Thus, it has been possible to use the extracted knowledge as basis for a new search strategy of text materials and increase the information available.

In this Poster we present the intelligent browsing system that operates on text materials, recovered and annotated by means of the TextPower technology.

### Bionotes

**Eugenio Picchi** graduated in Computer Science, at Pisa University, is Research Director at the Institute of Computational Linguistics (ILC) of the National Research Council. He is currently responsible for the research line "Computational models and tools for research in humanities, with a special focus on linguistic and literary disciplines and on lexicography". From 1972 to 1983 he was responsible of the Systems and Programming Division of the Linguistics Section of CNUCE (Pisa). Since 1983 he has been responsible of the Division "Methodologies and Tools for Lexicology and Computational Linguistics" of ILC. In the last few years he has been scientific director of national and international projects, among which: "International Network of Linguistics Data-Bases and Workstations" and "New Technologies for Language Engineering", within the Project "Natural Languages Processing"; ILC Unit of Research of Esprit Basic Research Actions "Aquilex - Acquisition of Lexical Knowledge for Natural Language Processing Systems", Action n. 3030; ILC Unit of Research of European Project MULTEXT "Multilingual Text Tools and Corpora". He has also been Technical Manager of CNR in the European Project "EUROSEARCH" for an European Federation of multilingual WEB browsers and Scientific Director of the Project "Italian-Arabic bilingual Corpora and Tools" within the CLUSTER "Computational Linguistics: monolingual and multilingual Researches", funding by Italian law 488/98. He has authored/co-authored a number of publications dealing with Computational Linguistics. [eugenio.picchi@ilc.cnr.it](mailto:eugenio.picchi@ilc.cnr.it)

**Eva Sassolini**, see page 109



## Tools and resources supporting the cultural tourism

Istituto di Linguistica Computazionale  
"Antonio Zampolli" CNR - Pisa

Sassolini Eva, Cinini Alessandra, Sbrulli Stefano, Picchi Eugenio

GL14: November 28, 2012

1

---

---

---

---

---

---

---

---

## Cultural Heritage domain

- Cultural Heritage is an open domain
  - large amount of information available from heterogeneous sources
  - hard to classify information through hierarchical criteria
  - approaches based on criteria of "semantic similarity" are desirable

GL14: November 28, 2012

2

---

---

---

---

---

---

---

---

## ICT for Cultural Heritage

- high diffusion on the Internet
- continuous development of information technologies
- advanced NLP technologies to discover domain-specific knowledge
- text annotation by means of terminology and named entities for:
  - {text} analysis
  - " browsing
  - " categorization

GL14: November 28, 2012

3

---

---

---

---

---

---

---

---

## Specific tools and Resources

- creation of linguistic resources
  - reference corpus
  - repository of terms pertaining to the cultural heritage domain
- extraction of relevant information from texts
  - named entities
  - single term
  - multi-words
- exploitation of the linguistic resources extracted
- text browsing

GL14: November 28, 2012

4

---

---

---

---

---

---

---

---

## Background/experience

- **Linguistic Miner**
  - system for the automatic extraction of linguistic knowledge from text
- **Text Power**
  - building of terminological resources
  - enrichment and annotation of textual material

GL14: November 28, 2012

5

---

---

---

---

---

---

---

---

## Text acquisition strategy (1)

Two phases for the automatic text acquisition:

1. *spidering* tools on institutional websites of historical, cultural, artistic and naturalistic interest
  - first Corpus building
  - generation/enrichment of linguistic resources

GL14: November 28, 2012

6

---

---

---

---

---

---

---

---

## Text acquisition strategy (2)

2. specialized crawlers that work on a bulk of text material available on-line

- using the extracted knowledge as basis for a new search strategy of text materials and increase the available information
- Reference (text) Corpus building

GL14: November 28, 2012

7

## projects & applications

### Online dissemination of the historical, artistic and landscaped regional heritage

- **Partner:** ILC, Basilicata APT (i.e. Agenzia di Promozione Territoriale della regione Basilicata)
- **Objectives:**
  - strategies implementation for the regional heritage promotion and dissemination
  - definition of a model for relevant information acquisition
  - automatic extraction of semantic information and terminology for the text categorization

**SMARTCITY:** new solutions for content engineering and ambient intelligence as support of cultural tourism

- **Partner:** ILC, consortium of companies (Space, Rigel Engineering, Meta)
- **Objectives:**
  - technologies for the cultural heritage preservation and enhancement;
  - methodologies and solutions to meet new demand of cultural spaces in particular for tourist purposes

GL14: November 28, 2012

8

**Faccette :**

abside anfiteatro

abside	anfrateito	animale marino	appio	area archeologica	asare
basilata	basinato	carattere sacro	cassa pirata	castello piro del balzo.	
catatonio	catatonica erutiva	causante alla	catena antri	catena di	chiesa
duce da trinità	colonia latina	complesso apicale	complesso termale	comunità abrica	corra
garibale	corso vitruvio	edificio pubblico	edificio termale	annuale	età repubblicana
et	et	et	et	et	et
funaria	fun letteraria	ipogeo	largo marcello	medusa	mon
necripo	opera pubblica	trane	parco archeologico	archeologia	proprietà
restaurato	restaurato	restaurato	restaurato	restaurato	struttura
arbitraria	struttura produttiva	testimonianza	archeologia	vasta	basilica
					venosa

GL14: November 28, 2012

0

## Purposes

- A very significant issue for e-Participation and e-Government is the need to know salient facts and features, hidden amidst a very large quantity of data.
- The collection of relevant information is useful in decision-making:
  - for a general user, to implement a service to the citizen ensuring transparency and the right to information;
  - creating a specialized database of valuable help for expert users

GL14: November 28, 2012

10

---

---

---

---

---

---

---

---

## Semantic networks for improved access to biomedical databases

**Eva Sassolini, Sebastiana Cucurullo, and Eugenio Picchi**

Istituto di Linguistica Computazionale, ILC, Italy

The proliferation of technologies has allowed an exponential growth of knowledge in the biomedical field. Most of the information available and the diversity of information sources are a strong constraint to the full exploitation of this knowledge.

The availability of systems for collecting and aggregating data and for semantic analysis has therefore become a priority. These systems allow a better exploitation of knowledge, in order to support both the public health, as well as economic growth based on knowledge.

The project “Unique Social Network for Innovation in Biomedical Tuscany” (SUBITO) funded by Regione Toscana, has allowed us to specialize an intelligent browsing system: “DBT-Faccette”, in order to facilitate the identification of links and potential scientific and technological synergies between researchers, government agencies and companies in Tuscany, which was the aim of the project.

The knowledge base Corpus contains abstracts of the scientific arguments extracted from biomedical bibliographic database PubMed for free web-users, including citations for biomedical articles from scientific journals. Starting from this data (that can be considered “gray information”), the system is able to relate semantically relevant information and suggest synergies between private and public institutions or subjects. The goal is to build a network of “knowledge” that can be exploited by an intelligent navigation system. This approach allows the automatic reorganization of content, based on the semantically relevant concepts extracted from the text, and the user is able to dynamically discover the concepts relevant for the domain. In this way it is possible to carry out search refinements through the interrelated concepts. Compared to traditional search engines, in addition to the intelligent navigation system semantically driven, “DBT-Faccette” also offers the ability to present the results obtained using various ways of ranking and clustering, allowing the explicitation of the most relevant information.

We have developed a “generalized” text enrichment technique, to point out all forms of knowledge identified in the text, through statistical and linguistic analysis technologies, free from predefined assumptions, structures and ontologies. All information extracted, are associated in a para-textual formalism: “text enrichment” of all known lexical, semantic, factual, named entities, terminologies. In this Poster we present “DBT-Faccette” and the results of some queries that are possible on SUBITO project contents.

### Bionotes

**Eva Sassolini** graduated in Computer Science, at Pisa University, is CTER (Research Collaborator) at the Institute of Computational Linguistics (ILC) of the National Research Council - Pisa. She is involved currently in several national and international projects. Research Collaborator in “TextPower” (TP) project, (new technology and approach to treatment and exploitation of texts) and before in the project “Corpus Bilingue Italiano-arabo” for linguistic tools and resources for bilingual Italian/Arabic corpora realization. Junior researcher ILC in the project LINGUISTIC MINER: linguistic Knowledge system for the Italian language; working contribution in “INTERA” (Integrated European Language Data Repository Area) project, for multilanguage terms extraction. Junior researcher ILC in the project: “Progetto Iraq: navigazione nei materiali testuali del museo virtuale di Baghdad in maniera contrastava tra le tre lingue previste dal progetto (italiano, arabo e inglese)”. Junior researcher ILC in the project 8: “Diffusione della cultura e valorizzazione del patrimonio letterario della lingua italiana e della lingua araba attraverso una diffusione telematica di banche di dati letterarie”. Collaboration with IMSS (Istituto e Museo di Storia della Scienza) for the realisation of web applications for the query on galileian texts. Junior researcher in the project: “Corpus Bilingue italiano-arabo”: in the framework of the comprehensive “Linguistica Computazionale: ricerche monolingui e multilingui”. Email: [eva.sassolini@ilc.cnr.it](mailto:eva.sassolini@ilc.cnr.it)

**Eugenio Picchi**, see page 104

**Semantic networks for improved access to biomedical databases**

Istituto di Linguistica Computazionale  
"Antonio Zampolli" CNR – Pisa

Sassolini Eva, Cucurullo Sebastiana, Picchi Eugenio

GL14, 28 November 2012

---

---

---

---

---

---

---

**Biomedical domain main features**

- large amount of information available
- several information sources
- problematic exploitation of knowledge due to unstructured textual format

GL14, 28 November 2012

---

---

---

---

---

---

---

**State of the art**

- risk of proliferation of disorganized information, difficult to find and to interconnect
- demand for tools that suggest a new modality of content access
- Internet and in particular social networking tools allow to create networks of people sharing common interests

GL14, 28 November 2012

---

---

---

---

---

---

---

## SUBITO (Unique Social Network for Biomedical Innovation in Tuscany)

the project allowed to specialize  
an intelligent browsing system:  
"DBT-Faccette"

GL14, 28 November 2012

4

---

---

---

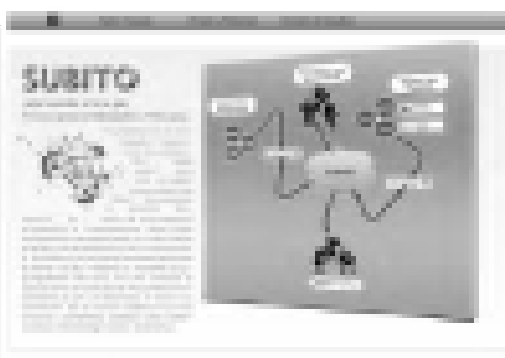
---

---

---

---

---



<http://subito-ifc.cnr.it/>

GL14, 28 November 2012

5

---

---

---

---

---

---

---

---

## Project partners



GL14, 28 November 2012

6

---

---

---

---

---

---

---

---

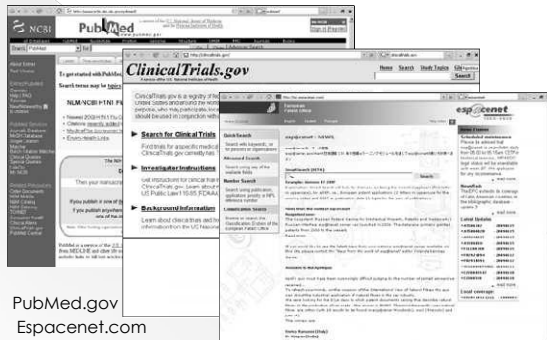
## OBJECTIVES

- (GENERAL) development of a website and a database for the collection and correlation of information on public-private subjects operating in the medical, biomedical and pharmaceutical sector within the Tuscan regional area
- (SPECIFIC) development of techniques for the extraction and classification of textual data in order to enable a more efficient browsing

GL14, 28 November 2012

7

## Sources of abstracts



- PubMed.gov
- Espacenet.com
- ClinicalTrials.gov

GL14, 28 November 2012

8

## Search approach

- Top-down:
  - the traditional search systems adopt a priori knowledge of the material, that can be combined to form user-defined keywords/queries
- Bottom-up:
  - the categorization systems allow the user to dynamically discover the concepts that are semantically relevant to a particular domain, and to carry out the search refinements on the basis of semantically related concepts, but virtually unknown

GL14, 28 November 2012

9



## DBT-Faccette

- labeling and textual analysis system as customization of the traditional faceted classification system
- "auto-adaptive" categorization that allows for automatic reorganization of content on the basis of semantically relevant concepts
- identification of the texts related to a given topic thanks to an automatically recognized terminological lexicon

GL14, 28 November 2012

10

---

---

---

---

---

---

---

## Search example



GL14, 28 November 2012

11

---

---

---

---

---

---

---

## Five Ways to Add Value to PhD Theses and Two Ways to do it Better

Joachim Schöpfel

Charles de Gaulle University Lille 3, France

### Problem/goal

Compared to articles published in academic journals, PhD theses are sometimes considered as second-level information, for at least two reasons. PhD theses do not undergo peer review, and they are produced in the early academic career. Yet, these documents often not only produce results from years of intensive research, together with a rich literature review (state of the art) but they are also validated by an academic committee and represent a growing part of open available content in institutional repositories (IR) and other open archives where they contribute to the impact and ranking of the institution.

More than the half of all IRs listed in the Directory of Open Access Repositories (OpenDOAR) contain electronic theses or dissertations (ETDs), and the Registry of Open Access Repositories (ROAR) lists more than 200 services with 100% ETDs. Mainly because institutions need a solution for the preservation and diffusion of PhD theses, some people even consider them as a kind of Trojan horse that pushes institutions to launch open repositories, even without full compliance or mandate from the scientific community.

All these projects and services have some similarities and common features, such as access to full text and OAI compliant interfaces. But more important are differences, with regards to metadata, policy, access restrictions, representativeness, file format, status, quality and related services.

Our communication deals with the last aspects. How can an institution add value to PhD theses in IRs and other open archives? What can be done to improve quality of content and service provision in an open environment, in order to increase impact, traffic and usage?

### Research method/procedure

The paper provides a state of the art of scientific and professional literature on PhD theses in institutional repositories and other open archives, with special attention to critical success factors, best practices and added value features. In particular, it reviews recent communications of international conferences on grey literature and electronic theses organized by the Grey Literature Network Service (GreyNet), of symposia organized by the Networked Digital Library of Theses and Dissertations (NDLTD) and of related events (2007-2012). The paper draws also on our own work on PhD and Master theses, on grey literature in open archives and on development, usage and emerging standards of institutional repositories. It will include benchmark examples from academic libraries, information services and repositories.

### Costs related to the project

20 days = approx. 4,000.00 EUR (only salary, without travel costs).

### Anticipated results

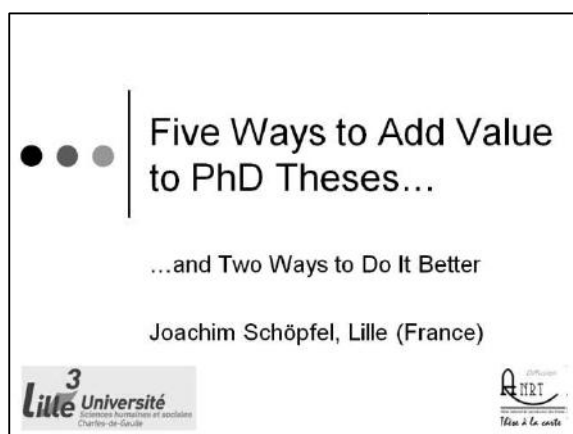
An overview on some models of PhD theses related services. Five recommendations how to add value to them. Two further perspectives on the preservation and dissemination of PhD theses. The communication is meant to be helpful for service provider of institutional repositories that have to take decisions on PhD theses.

## Bibliography

- N. Avdeeva (2010). 'Innovative services for libraries through the Virtual Reading Rooms of the Digital Dissertation Library, Russian State Library'. *IFLA Journal* **36**(2):138-144.
- E. Bester (2010). 'Les services pour les archives ouvertes : de la référence à l'expertise'. *Documentaliste - Sciences de l'Information* **47**(4):4-15.
- J. Ducloy, et al. (2006). 'Metadata towards an e-research cyberinfrastructure: the case of French PhD theses'. In *DCMI '06: Proceedings of the 2006 international conference on Dublin Core and Metadata Applications*, pp. 133-148. Dublin Core Metadata Initiative.
- E. G. Park & M. Richard (2011). 'Metadata assessment in e-theses and dissertations of Canadian institutional repositories'. *The Electronic Library* **29**(3):394-407.
- A. Sale (2006). 'The impact of mandatory policies on ETD acquisition'. *D-Lib Magazine* **12**(4).
- J. Schöpfel, et al. (2011). 'Open is not enough. A case study on grey literature in an OAI environment'. In *Thirteenth International Conference on Grey Literature: The Grey Circuit. From Social Networking to Wealth Creation*. Washington, DC, 5-6 December 2011.
- C. Stock (2007). 'Open access to full text and ETDs in Europe: improving accessibility through the choice of language?'. In *Ninth International Conference on Grey Literature: Grey Foundations in Information Landscape*, Antwerp, 10-11 December 2007.

## Bionote

**Joachim Schöpfel** is Head of the Department of Information and Library Sciences at the Charles de Gaulle University of Lille 3 and Researcher at the GERiiCO laboratory. He is interested in scientific information, academic publishing, open repositories, GL and usage statistics. He is a member of GreyNet and euroCRIS. He is also the Director of the National Digitization Centre for PhD Theses (ANRT) in Lille, France. [joachim.schopfel@univ-lille3.fr](mailto:joachim.schopfel@univ-lille3.fr)





## Methodology

- o State of the art
  - NDLTD/ETD 2007-2012
  - USETDA 2010 & 2011
  - GL (in OpenGrey)
- o Studies on open access and grey literature
- o Case studies

GL14 Rome 2012

---

---

---

---

---

---

---



## Five ways to add value to PhD theses

- o Quality
- o Metadata
- o Format
- o Interoperability
- o Services

GL14 Rome 2012

---

---

---

---

---

---

---



## Quality

- o *"The content (...) is the most important factor that has been cited by researchers to show the success of a repository"* (Macha & Jager 2011)
  - Policy, standards
  - Only validated version
  - Review system
  - Requirement of minor or major revisions
  - Minimal requirements for acceptance (OA, fulltext, research theses, metadata...)
  - Born-digital and older repurposed digital materials
  - Retrospective digitization of selected sources
  - Unrestricted access
  - Copyright cleared

GL14 Rome 2012

---

---

---

---

---

---

---

## Metadata

- *"Considerable variations in practice exist (...) Many of these variations involve local qualifiers and/or the use of metadata elements to record information that does not correspond to the prescribed meaning of the elements"* (Park & Richard 2011)
  - Standards, good practice
    - OAI DC simple (if networking, cf. DART Europe)
    - NDLTD standards for DC/MARC (ETD-MS)
    - ProQuest schema (UMI XML DTD)
    - TEF (France)
    - EVSKP-MS (Czech Republic)
  - Especially for small and medium sized institutions
  - Bibliographic description for searchability
  - Value added metadata for scientometrics
  - Metadata for print on demand (book publishing)
  - Metadata must change with new policies (Heyse 2007)
  - Easy conversion to other formats

GL14 Rome 2012

## Format

- *"In the sciences theses should be regarded not simply as objects destined for preservation and archival status, but rather as unique resources containing potentially valuable data that must be made extractable and re-usable"* (Morgan et al. 2008)
  - Fulltext and not just metadata
  - PDF
  - LaTeX → XML
  - Postscript → RDF
  - MS Office Word DOC
  - MS Office Open XML DOCX
    - Format processing
    - OCR
    - Semantic querying

GL14 Rome 2012

## Interoperability

- *"Demonstrate research efforts of the state... Central finding tool for research"* (Dowling & Steans 2011)
  - OAI-PMH
  - Policy, statement, commitment, collective will
  - National network
  - Statewide architecture
  - International network
  - Single hub for research
  - Discoverability of research

GL14 Rome 2012

● ● ●

## Services

- *"User-Centered Process of Service Development Prioritization"* (Halbert 2007)
  - Flexibility and capacity for rapid adaptation
  - Interactivity, discussion forum, mailing list, comments
  - Discovery tools, federated search
  - User-friendly and reliable software
  - Statistics and metrics, citations
  - Video with presentation of thesis
  - XML export
  - Print on demand
  - Information literacy education
  - © or CC licensing
  - Preservation (multiple copies)
  - Extended (distant) learning environment

GL14 Rome 2012

---

---

---

---

---

---

---

---

● ● ●

## Two ways to do better

- CRIS
- Data

GL14 Rome 2012

---

---

---

---

---

---

---

---

● ● ●

## CRIS

- *"A well managed repository can simplify the task for the researchers and their respective departments, their university administrators and (...) their assessors"* (Hey et al 2006)
  - ETDs as important part of IR integrated in national infrastructure
  - Assignment of unique identifier
  - OrgID and PersID
  - Analysis of research trends

GL14 Rome 2012

---

---

---

---

---

---

---

---

● ● ●

## Data

- *"It may be possible to enhance both the research options and the richness of access. "Deep access" to the research could include data, visual and aural representations, links to supporting material such as notebooks or collaborative activity."* (White 2007)
  - Supplemental content
  - Video and multimedia (avi)
  - Images
  - Spreadsheets
  - Music (mp3)
  - Attn with copyright and formats

GL14 Rome 2012

---

---

---

---

---

---

---

---

● ● ●

## Examples

- Emory ETD Program (Halbert 2007)
- Peruvian Network of Digital Theses (Huarato 2008)
- Humboldt University edoc server (Schirnbacher 2009)
- Butler University (Tamminga 2010)
- EthOS (Gould & Rosie 2012)
- DART E-theses @ UCL (Moyle 2009)
- ProQuest metadata (Marsh & McLean 2008)
- Format (Morgan et al 2008)
- Quality (Lewandowska 2010)
- CRIS Soton (Hey et al 2006)

GL14 Rome 2012

---

---

---

---

---

---

---

---

● ● ●

## Bibliography and contact

<http://www.citeulike.org/user/Schopfel/tag/gl14theses>

Email [joachim.schopfel@univ-lille3.fr](mailto:joachim.schopfel@univ-lille3.fr)

ANRT <http://www.diffusiontheses.fr/>

GL14 Rome 2012

---

---

---

---

---

---

---

---

## An Environment Supporting the Production of Live Research Objects

**Massimiliano Assante, Leonardo Candela, Donatella Castelli, and Pasquale Pagano**

Istituto di Scienza e Tecnologie dell'Informazione – CNR, Italy

Scientific research is rapidly evolving in all fields, it is multidisciplinary, networked and driven by new patterns, e.g. data-intensive sciences. In this complex scenario scientific communication must go well beyond traditional scholarly publications. In particular, it requires accessing all the elements exploited and developed during the scientific workflow to achieve a result, e.g. datasets, analysis tools, and methods. This wide corpus of primarily grey elements is at the moment mostly unavailable and, even when they are available, they are not linked to the scientific result. This makes difficult to completely understand the result and validate it. To overcome this limitation “live research objects” are emerging as replacement of traditional “documents” in the scientific communication. These are multi-media and multi-part objects that aggregate all “pieces” contributing to a research result. Such elements, which may range from binary files to compound objects including maps, time series, and tabular data, are generally structured according to well-established templates and produced according to user-defined workflows.

In this paper, we present an environment for the production of live research objects. Such environment includes: (i) a workspace where users can organise and share in a file-system-like environment items of very different types (from binary files to compound objects); (ii) an editing framework where users can define the structure of a live research object (a template indicating sections, layout, active elements) and compile objects compliant with one of the defined templates by entering content or taking it from the workspace via drag & drop; and (iii) a workflow engine where users can define the workflow governing the production of a live object by specifying the phases and the relative responsible actors(s).

Supporting these new live research objects outgrows the capacity of traditional technological approaches based on locally specialized data management facilities. The paper also briefly discusses how the above facilities have been realised by exploiting the capabilities offered by an Hybrid Data Infrastructure.

The exploitation of the above described supporting environment is discussed by exemplifying production of research objects aiming at estimating the probability of distribution of biological species in a global scale. Such objects comprise some descriptive text, data on the species gathered from authoritative data sources, environmental data reporting on ecological elements, the algorithm(s) aiming at estimating the probability of occurrence of a species in a given area, and the maps resulting from the algorithm(s).

The environment described has been designed and implemented in the context of two successive EU projects: D4Science-II ([www.d4science.eu](http://www.d4science.eu)) and iMarine ([www.i-marine.eu](http://www.i-marine.eu)). It is now available as a gCube Apps in the D4Science e-Infrastructure (<https://portal.i-marine.d4science.org/>).

### Bionotes

---

**Massimiliano Assante** holds a master degree (M.Sc.) on Information Technologies received by the University of Pisa, where he is undertaking a Ph.D. on Information Engineering. He is research staff at the Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo" (ISTI), an institute of the Italian National Research Council (CNR). He joined ISTI in 2007 and is currently member of the iMarine EU Project and EUBrazilOpenBio Project. In the past he has been member of D4Science II, D4Science, DILIGENT and DRIVER European Projects. His research interests include Data Infrastructures, Next Generation Digital Libraries and Information Systems. Email: [assimiliano.assante@isti.cnr.it](mailto:assimiliano.assante@isti.cnr.it)



## Bionotes (continued)

**Leonardo Candela** is a researcher at Networked Multimedia Information Systems (NeMIS) Laboratory of the Italian National Research Council - Institute of Information Science and Technologies (CNR - ISTI). He graduated in Computer Science in 2001 at University of Pisa and completed a PhD in Information Engineering in 2006 at University of Pisa. He joined the NeMIS Laboratory in 2001 and was involved in various EU-funded projects including CYCLADES, Open Archives Forum, DILIGENT, DRIVER, DELOS, D4Science, D4Science-II and DL.org. He was a member of the DELOS Reference Model Technical Committee and of the OAI-ORE Liaison Group. He is currently involved in the iMarine and EUBrazilOpenBio projects. His research interests include Digital Library [Management] Systems and Architectures, Digital Libraries Models, Data Infrastructures. Email: [leonardo.candela@isti.cnr.it](mailto:leonardo.candela@isti.cnr.it)

**Donatella Castelli** is a Senior Researcher working at the "Istituto di Scienza e Tecnologie della Informazione" of the Italian National Research Council (CNR-ISTI). She has been the principal investigator of several European and National funded projects on digital libraries and data e-Infrastructure acquiring considerable experience in these domains. Currently she is acting as scientific director of the iMarine ([www.i-marine.eu](http://www.i-marine.eu)) and EUBrazilOpenBio ([www.eubrazilopenbio.eu](http://www.eubrazilopenbio.eu)) EU projects. She is also technical coordinator of the OpenAIREPlus EU projects which aim at developing and operating the Open Access e-infrastructure supporting the scientific communication produced in the context of EU funded projects. Email: [donatella.castelli@isti.cnr.it](mailto:donatella.castelli@isti.cnr.it)

**Pasquale Pagano** is a Senior Researcher at the Networked Multimedia Information Systems Laboratory of the "Istituto di Scienza e Tecnologie della Informazione A. Faedo" (ISTI) of the Italian National Research Council (CNR). He received my M.Sc. in Information Systems Technologies from the Department of Computer Science of the University of Pisa (1998), and the Ph.D. degree in Information Engineering from the Department of Information Engineering: Electronics, Information Theory, Telecommunications of the same university (2006). The aim of his research is the study and experimentation of models, methodologies and techniques for the design and development of distributed virtual research environments (VREs) which require the handling of heterogeneous resources provided by Grid and Cloud based e-Infrastructures. Pasquale has a strong background on distributed architectures. He participated to the design of the most relevant distributed systems and e-Infrastructure enabling middleware developed by ISTI - CNR. Pasquale is currently the Technical Director of the Data e-Infrastructure Initiative for Fisheries Management and Conservation of Marine Living Resources (iMarine) and member of VENUS-C European project. He is also involved in the GRDI2020 expert working group a he serves EUBrazilOpenBio initiative as consultant. In the past, he has been involved in the D4Science-II, D4Science, Diligent, DRIVER, DRIVER II, BELIEF, BELIEF II, Scholnet, Cyclades, and ARCA European projects. Email: [pasquale.pagano@isti.cnr.it](mailto:pasquale.pagano@isti.cnr.it)



## Scientific research today



- Scientific research is rapidly **evolving** in all fields, it is multidisciplinary, networked and driven by new patterns, e.g. data-intensive sciences
- Scientific communication should go well beyond **traditional** scholarly publications
- Requires access to **all the elements** exploited and developed to achieve a **result**
  - datasets, analysis tools, methods, etc.



GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

---

---

---

---

---

---

---

---

## Understanding Scientific Research



*"Publishing data in a reusable form to support findings must be mandatory" [Science as an Open Enterprise, The Royal Society]*

- The elements exploited and developed during the scientific workflow to achieve a result (primarily grey elements) are **not available** or **not linked** to the scientific result
  - from binary files to compound objects including images, maps, tabular data, time series, etc.

- It makes difficult to completely **understand** the results
- It makes difficult to **validate** the results



GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

---

---

---

---

---

---

---

---

## Live research objects



- Emerging as **replacement** of traditional "documents" in the scientific communication
- multi-media and multi-part "documents" that **aggregate** all the "pieces" that contribute to a research result
- Supporting them has gone **beyond** the capacity of traditional technological approaches based on locally specialized data management facilities



GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

---

---

---

---

---

---

---

---

## Production of live research objects

Our environment includes:

### ➤ A Virtual Workspace

- to handle a rich array of objects from binary files to compound information objects

### ➤ An Editing framework

- to define the structure of a live research object
- to simplify live research objects compilation

### ➤ A Workflow Engine

- to define the workflow governing research objects production
- to specify phases, roles and users

GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome



## Virtual Workspace

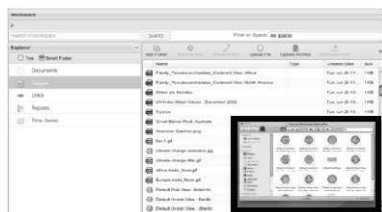
- A working area where users can organise and share very different items (from binary files to compound objects) in a file-system-like environment

- e.g. tabular data, species distribution maps, time series

### ➤ Sharing

### ➤ Smart Folders

### ➤ Desktop Access

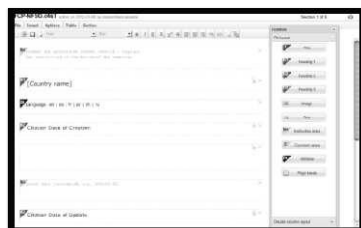


GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

## Editing Framework

- Template definition phase, to create a document template indicating sections, layout, attributes and active elements

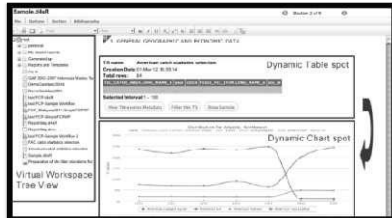
- define the structure of a live research object
- component oriented approach (static & dynamic)



GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

## Editing Framework

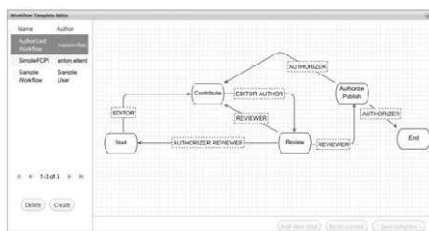
- Compiling phase: to compile a live research object, that complies with one of the defined templates by entering content or taking it from the workspace
  - complete or instantiate the dynamic components



GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

## Workflow Engine – defining a workflow

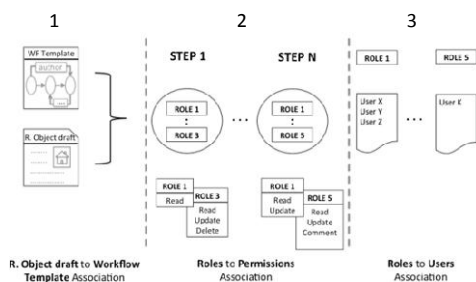
- Allows to work collaboratively to the creation of a research object
- Define the workflow governing the production of a research object




GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

## Workflow Engine – associating Research Objects

- Specify the steps and the relative responsible actor(s)
  - made of 3 Phases



GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome



## Monitoring and Accessing Research Object workflows

- Workflow owner can continuously **monitor** the users activities involved in a workflow
  - manual routing and automatic routing are supported
- Sharing** of Research Object workflows through virtual workspace
  - Special folder: My R. Object Workflows
  - Concurrent access guaranteed by a locking mechanism

GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

---

---

---


---

---

---

---

---



## Demonstration

- Exploited in a number of diverse scenarios, e.g. the production of Live Research Objects aimed at estimating the probability of biological species distribution in a global scale, containing
  - some descriptive **text**
  - data on the species gathered from authoritative **data sources**
  - environmental data** reporting on ecological elements
  - algorithms** aimed at estimating the probability of the occurrence of a species in a given area
  - images of **maps** resulting from the algorithm(s)

GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

---

---

---


---

---

---

---

---



## Conclusion

- A comprehensive framework supporting the entire lifecycle of Live Research Objects production and management
- It has been designed and implemented in the context of two successive EU projects:
  - D4Science-II ([www.d4science.eu](http://www.d4science.eu))
  - iMarine ([www.i-marine.eu](http://www.i-marine.eu))
- Available as a Web App in the D4Science e-Infrastructure (<https://portal.d4science.org/>)

GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome

---

---

---

---

---

---

---

---

## Creating and Assessing a Subject-based Blog for Current Awareness within a Cancer Care Environment

Yongtao Lin and Marcus Vaska

Health Information Network Calgary, University of Calgary, Canada

### Introduction

The Health Information Network Calgary (HINC) is comprised of a group of libraries providing information services and resources to urban and rural sites in the Calgary Zone of Alberta Health Services. Establishing a current awareness service is a necessity in any discipline, especially in health care. Web 2.0 and social networks have transformed how health care professionals and researchers create knowledge, access information, collaborate, and disseminate research.

One of the earliest forms of social media, blogging has taken the world by storm (1). Although there is a wealth of literature on the use of blogs in providing current awareness services for libraries, there is a pronounced gap on how blogs are assessed or evaluated, especially for information alert purposes (2).

### Background

Clients within the HINC subscribe to e-mail alerts and RSS feeds, a trend particularly evident within the Cancer Care environment where a number of researchers have already implemented feed readers to remain aware of current literature. However, they often comment on challenges associated not only with maintaining alerts and managing RSS feeds, but also in selecting and creating alerts for unpublished materials. The need for a librarian-facilitated current awareness strategy became more and more apparent. The literature reviewed addressed the value of an alert, namely to indicate a gap in the participant's knowledge, rather than to deliver content the librarians may have perceived as useful (3). The authors saw the creation of a subject-based blog as an opportunity to disseminate current awareness "grey" information to this specific research community.

### Method/Anticipated Results

The Grey Horizon Blog was created in April 2012 using Blogger. The selection and re-aggregation of information involves ongoing assessment of user needs and continuous work on the Blog. A weekly global email-digest listing of the postings will be distributed two months after the launch.

Several metrics will be employed in October 2012 to evaluate the Blog. Blogger itself tracks the number of page-views over time. Google Analytics was set up as it tracks additional information on access and use of the Blog. As clients may be using feed readers to read blog entries and may thus not visit the Blog at all, Feedburner has also been incorporated to track the number of times that the blog RSS is accessed, as well as calculating the number of subscribers.

A post-survey will be conducted in six months to complement the web statistics data. The additional feedback and comments will help us determine whether the Blog has successfully created an easy platform for users to keep current with unpublished literature, the type of resources found most important, and whether the amount of time spent maintaining the Blog met expectations.

### Discussion

It is anticipated that this case study will portray how to successfully plan a subject-based blog to meet users' current awareness information needs in grey literature. Further efforts will focus on targeting the blog to the topic areas in grey literature where users feel more information is needed. The findings from this assessment will direct us to potential marketing opportunities and changing technology that haven't been fully utilized in our Grey Horizon Blog.

## References

- (1) Graham DL. Social Media and Oncology: Opportunity with Risk - ASCO Available at: [http://www.asco.org/asco2/Education+&+Training/Educational+Book?&vmview=edbk\\_detail\\_view&confID=102&abstractID=1054](http://www.asco.org/asco2/Education+&+Training/Educational+Book?&vmview=edbk_detail_view&confID=102&abstractID=1054). Accessed 4/12/2012, 2012.
- (2) Blair J, Level AV. Creating and evaluating a subject-based blog: planning, implementation, and assessment. *Reference Services Review* 2008;36(2):156-166.
- (3) Attfield S, Blandford A. Conceptual misfits in e-mail-based current-awareness interaction. *Journal of Documentation* 2011;67(1):33-55.

## Bionotes

**Yongtao Lin** has been working as a health information network librarian at the Tom Baker Cancer Knowledge Centre in Calgary Alberta since 2008. She provides library services to support health care professionals in their evidence-based practice. The library is part of a provincial patient-centered education strategy supporting cancer patients and their families. She was a hospital librarian in rural Nova Scotia for a few years before moving to the University of Calgary. Her prior experience as an instructor has led her to integrate education into various aspects of library programs. Yongtao is interested in the impact of grey literature in health care and a strong believer in evidence-based practice. Yongtao was awarded the Canadian Hospital Librarian of the Year in 2011. Email: [yolin@ucalgary.ca](mailto:yolin@ucalgary.ca)

**Marcus Vaska** is a librarian with the Health Information Network 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. Marcus' current interests focus on educational techniques aimed at creating greater awareness and bringing grey literature to the forefront in the medical community. Email: [mmvaska@ucalgary.ca](mailto:mmvaska@ucalgary.ca)

## Creating and Assessing a Subject-Based Blog for Current Awareness within a Cancer Care Environment

Yongtao Lin, MLIS [yolin@ucalgary.ca](mailto:yolin@ucalgary.ca)  
 Marcus Vaska, MLIS [mmvaska@ucalgary.ca](mailto:mmvaska@ucalgary.ca)

Health Information Network Calgary,  
 University of Calgary

GL14, Rome, November 30, 2012




---

---

---

---

---

---

---

---

## Social Media and Grey Literature Are Healthcare Professionals Ready?



### Purpose/Reason

- Engagement, development, support
- Knowledge dissemination
- Social norms
- Marketing
- Convenient and easy access to information (one place)

### Research Findings

- Over 50% of physicians in Europe and 41% in the U.S. regularly engage/post in blogs (Graham, 2011)
- 80% of medical organizations surveyed plan to implement blogs in their information pursuits; 50% currently do so (Chu, 2011)

---

---

---

---

---

---

---

---

## Current Awareness and Health Information

- **Open Access** movement
- **Sharing/promoting** of information
- **Improving** teaching and learning pursuits
- **Understanding** needs of users
- **Filtering & re-organizing** information
- **Aggregating** information




---

---

---

---

---

---

---


---



## What Have We Done in Integrating Social Media & Current Awareness?

Home >> Have a Question? Get Answers Fast via Chat!

Have a Question? Get Answers Fast via Chat!



Check out our YouTube channel for HHC tutorials and some:

- HHC An Introduction
- HHC How to Register for a Library Account
- HHC How to Renew Your Books Online

### Social Media in Library Services

- IM Chat
- Twitter
- You Tube Channel

### Current Awareness in Cancer Care

- In-person training
- Virtual presence
- Subject expertise

---

---

---

---

---

---

---

---

## Grey Horizon Process/Methodology

- Blogger as the platform
- Scope and breadth of content
- Selection of sources
- Frequency of postings
- Organization of information



5

---

---

---

---

---

---

---

---

## Grey Horizon Execution & Performance

- Soft launch
- Bi-weekly digest
- Promotion
- Project roll-out

Grey Horizon: Bi-weekly Digest

- New safety information: interaction of proton pump inhibitors with methotrexate
- Wide variation in immunotherapy rates across Canada
- Genomic Cancer Center researchers discover mechanism
- SAHRI: First large-scale trial of whole-genome cancer testing for clinical decision-making
- Maintaining treatment with MGMT? Following standard first-line treatment protocols

How to access full-text articles

Published articles are available to staff of Alberta Health Services Cancer Care sites at Tom Baker Cancer Centre, Holy Cross Site, Lethbridge Cancer Care, and Medicine Hat Cancer Centre through the Health Information Network (HIN).

To access full-text articles from hundreds of medical journals licensed through HIN, visit our website at <http://hin.ahs.ca/gazette>, select Find a Specific Article, AHS Cancer Care, and search for the journal title.

---

---

---

---

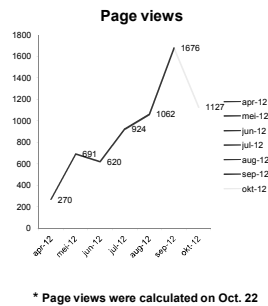
---

---

---

---

## Blog Analysis - *Blogger*



## Most popular posts

- **News** about a specific research conference (84)
- A cancer **drug update** from Health Canada (56)
- A **report** on cancer health services (53)
- A **clinical guideline** from NICE (51)
- **Statistics** on cancer incidence, survival and risk factors (46)

---

---

---

---

---

---

## Blog Analysis – Google Analytics



- 1785 “people” visited the site (identified by IP addresses)
- Average visit duration: 4 minutes 14 seconds

	1990	1995	2000	2005	2010
1990	1990	1995	2000	2005	2010
1995	1990	1995	2000	2005	2010
2000	1990	1995	2000	2005	2010
2005	1990	1995	2000	2005	2010
2010	1990	1995	2000	2005	2010
2015	1990	1995	2000	2005	2010
2020	1990	1995	2000	2005	2010
2025	1990	1995	2000	2005	2010
2030	1990	1995	2000	2005	2010
2035	1990	1995	2000	2005	2010
2040	1990	1995	2000	2005	2010
2045	1990	1995	2000	2005	2010
2050	1990	1995	2000	2005	2010
2055	1990	1995	2000	2005	2010
2060	1990	1995	2000	2005	2010
2065	1990	1995	2000	2005	2010
2070	1990	1995	2000	2005	2010
2075	1990	1995	2000	2005	2010
2080	1990	1995	2000	2005	2010
2085	1990	1995	2000	2005	2010
2090	1990	1995	2000	2005	2010
2095	1990	1995	2000	2005	2010
2100	1990	1995	2000	2005	2010

---

---

---

---

---

---

## Blog Analysis - *Feedburner*

- 17 subscribers in total
- 12 active subscribers on average
- 35% REACH on average every week
- Additional page views calculated

[illegible]

---

---

---

---

---

---

## Tweets & Staff Log: Connecting the Internal to the External




---

---

---

---

---

---

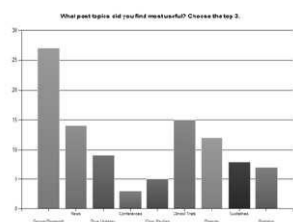
---

---

## Post-Survey

Four sections:

- Grey literature awareness and current awareness (CA) efforts
- Experience with the Blog
- Experience with bi-weekly digests
- Additional comments about CA services




---

---

---

---

---

---

---

---

## Discussion/Future Directions



- Subject-based blogs work well with CA service
- Blogs work well for grey literature
- Enhanced collaboration and information sharing
- Constant and various methods of marketing/promoting
- Thorough project planning (staff time, workload and user expectations)
- Expansion of other subject blogs

---

---

---

---

---

---

---

---

# 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

## Centralised National Corpus of Electronic Theses and Dissertations

**Július Kravjar and Marta Dušková**

Slovak Centre of Scientific and Technical information, CVTISR, Slovakia

ETDs are the significant source of grey literature and not only for the academic community. Slovakia has made a big step forward by implementation of Centralised National Corpus of Electronic Theses and Dissertations in 2010. National ETD corpus consists of bachelor's, master's, dissertation and habilitation theses. This implementation was coupled with the same time implementation of National Plagiarism Detection System (aka originality check or anti-plagiarism system). Both systems are obliged to use by all higher education institutions operating under the Slovak legal order. The new theses and dissertations incoming into the national ETD corpus are compared to this corpus and to selected internet resources. The higher education institutions pay no fee for the service, the system acquisition costs were covered by the Ministry of Education, Science, Research and Sport, the operating costs are also paid by the Ministry. The formation of both systems and the first two years of its existence is analyzed.

The first signs of activities towards ETD in Slovakia were recorded on the threshold of the Millennium. March 2004 was to become a significant milestone: sixteen academic libraries of twelve Slovak universities decided to solve the ETD.SK project: "Building Digital Academic Libraries - Collecting and Providing Access to Full Texts of Slovak University Publications". The ETD.SK project marked the beginnings of cooperation on a national level in this area, with the effort to follow up the international ETD activities. Unfortunately, the project was not sufficiently implemented due to the lack of financial and personnel resources, but mainly because of the lack of legislative support.

The ICT and internet penetration, low copyright awareness and the rapid growth in the number of higher education institutions and students in our country contributed to the expansion of plagiarism. There was also an inherent lack of systemic action, which would act as a barrier for its future growth. The establishment of a nationwide electronic theses and dissertation repository and their originality check was considered as a perspective solution.

A significant step in this matter was made in 2008: the Ministry of Education decided to implement a comprehensive nationwide solution for the collection and processing of theses and dissertations produced at Slovak higher education institutions. The goal: creation of the national theses and dissertations repository, increase in the quality of theses by their originality check, copyright and intellectual rights protection. In 2009, the Higher Education Act was amended and the most relevant change was this: Before the defence of the thesis, the higher education institution forwards the thesis in the electronic form to the Central Repository of Theses and Dissertations (CRTD) and the originality check is performed.

During 2009, the CRTD was built and the whole system with the originality check became reality at the end of April 2010. The existence of such a system has a preventive effect, and not just in the student community. CRTD is now publicly available source of grey literature.

Yearly increase in CRTD is about 80 thousand items of bachelor's, master's, dissertation and habilitation theses.

### Bionotes

---

**Július Kravjar** graduated in Mathematics at the Comenius University of Bratislava and later in Informatics. He is currently responsible for the „Central Repository of Theses and Dissertations“ and „Plagiarism Detection System for Slovak Academic and Research Institutions“ projects at the Slovak Centre of Scientific and Technical Information (SCSTI). Both nationwide systems are in the real operation from April 2010. He also participates as a team member at three other national projects: National Information System Promoting Research and Development in Slovakia - Access to Electronic Information Resources, Infrastructure for R&D - Data Centre for Research and Development, and National Infrastructure for Technology Transfer in Slovakia. Prior to joining the SCSTI, he held several positions in software development and software and ICT services marketing in the private sector. Email: [julius.kravjar@cvtisr.sk](mailto:julius.kravjar@cvtisr.sk)

## Bionotes (continued)

---

**Marta Dušková** studied library and information science at Comenius University in Bratislava (Slovakia). Since July 2010 she works in the Slovak Centre of Scientific and Technical Information in Bratislava (Slovakia) in Publication Evaluation Department. She deals with grey literature and The Central Registry of Publication Activity. She coordinates activities associated with obtaining and making grey literature available and cooperates with the processing and verification data publications included in The Central Registry of Publication Activity. She also participates as a team member at three national projects: National Information System Promoting Research and Development in Slovakia - Access to Electronic Information Resources, Infrastructure for R&D - Data Centre for Research and Development, and National Infrastructure for Technology Transfer in Slovakia. From 2012 she is studying PhD study at Comenius University in Bratislava (Slovakia) with the theme: Knowledge communication with support of grey literature.

Email: [marta.duskova@cvtisr.sk](mailto:marta.duskova@cvtisr.sk)

**14<sup>th</sup> INTERNATIONAL CONFERENCE ON GL**

**» CENTRALIZED NATIONAL CORPUS OF  
ELECTRONIC THESES AND DISSERTATIONS «**

**Július KRAVJAR, Marta DUŠKOVÁ**



**Slovak Centre of Scientific and Technical Information**

**ROMA, 29.-30.11.2012**

**AGENDA**

- ☛ *Background*
- ☛ *Pillars*
- ☛ *The Uniqueness*
- ☛ *Two Years After*

**BACKGROUND**

- ☛ *Low copyright and IPR awareness*
  - ☛ *HEI / students / internet penetration*
    - ☛ *public*                      20
    - ☛ *state*                         3
    - ☛ *private*                    12
    - ☛ *private foreign*        4
- |             |             |              |               |
|-------------|-------------|--------------|---------------|
| <b>2011</b> | <b>39 ↑</b> | <b>250 ↑</b> | <b>79.2 ↑</b> |
| <b>1989</b> | <b>13 ↓</b> | <b>63 ↓</b>  | <b>↓</b>      |

### BACKGROUND

- ☛ *Direct experience with cheating students + recurring discussions: T&D collection , the plagiarism matters*
- ☛ *The lack of measures, procedures, ... used at HEI to fight plagiarism*
- ☛ *The significant initiative of Slovak Rector's Conference*

### BACKGROUND

*2008: EduMin decision concerning HEI*

- ☛ *Two birds (systems) with one stone:*
  - ☛ *National repository: collection and processing of bachelor's, master's, dissertation & habilitation theses*
  - ☛ *Plagiarism detection system on a national level*
- ☛ *Launch date: 2010 – obligatory for HEI; costs paid by EduMin*

### PILLARS

- 1. The Higher Education Act*
- 2. The Methodological Guidelines, the intra-HEI directives*
- 3. Prepared staff, HW, SW infrastructure + the skills and the will*
- 4. Cooperating systems of different kind and origin: AIS, LRTD, CRTD, PDS*
- 5. EduMin support*



## THE UNIQUENESS

- *CRTD + PDS is obligatory for all HEI operating under the Slovak legal order*
- *To date, we have no information about a similar model in real operation*
- *The CRTD + PDS implementation on a national level into everyday practice is very likely a unique solution in Europe and perhaps in the world*

---

---

---

---

---

---

---

## TWO YEARS AFTER

- *CRTD + PDS = breakthrough in plagiarism fight on a national level*
- *Preventive effect*
- ↑ *Copyright & IPR protection awareness*
- ↑ *The quality of theses by their originality check*
- ↑ *The work of students with literature, internet, and quotes*

---

---

---

---

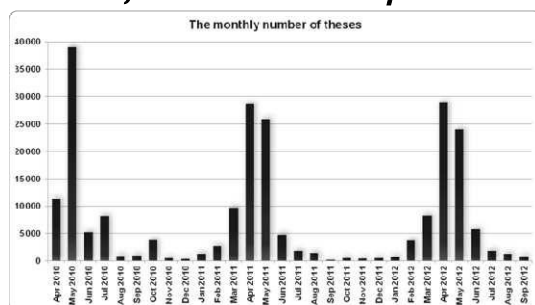
---

---

---

## TWO YEARS AFTER

**223,295 end of Sep 2012**




---

---

---

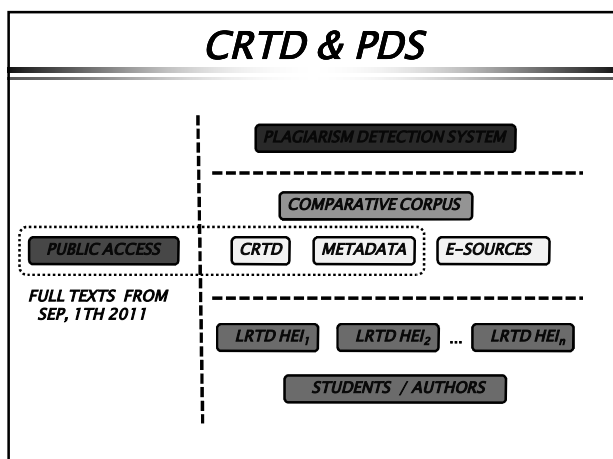
---

---

---

---

## CRTD & PDS




---

---

---

---

---

---

---

---

## RELATED PROJECTS

- **CRTD + PDS are running now in the Data Centre for R&D (DC R&D)**
- **National projects**
  - **DC R&D – The infrastructure for Research and Development**
  - **NITT SK – National infrastructure for supporting technology transfer in Slovakia**

---

---

---

---

---

---

---

---

## THE FUTURE STARTS TODAY



*Yesterday's the past, tomorrow's the future, but today is a gift.  
That's why it's called the present.*  
B. Keane

**THANK YOU FOR YOUR ATTENTION**

---

---

---

---

---

---

---

---

## A funder repository of heterogeneous grey literature material with advanced user interface and presentation features

**Ioanna-Ourania Stathopoulou, Nikos Houssos, Panagiotis Stathopoulos, Despina Hardouveli, Alexandra Roubani, Ioanna Sarantopoulou, Alexandros Soumplis, Chrysostomos Nanakos**  
National Documentation Centre; National Hellenic Research Foundation, Greece

The present contribution concerns the development of a funder repository aiming at the dissemination, reuse and preservation of mainly grey literature material of diverse types. This material which was produced under the auspices of large scale (multi-billion Euros) funding programmes of the Hellenic Ministry of Education (co-financed by the European Union). The project involved the handling of a wide range of content, like (among others) studies, reports, educational material, videos, theses, material from a range of conferences/events. The project has been successfully completed and the system is publicly available since spring 2011 at <http://repository.edulll.gr>. In this repository creation use case, technical enhancements were used to provide the means to organise, process and import to the repository a wide range of heterogeneous material. Special facilities for the support of these workflows were included. The metadata schema used is an application profile using elements for Qualified Dublin core, LOM and PREMIS.

A major part of the work concerned mechanisms to enhance presentation of the digital material and better navigation to the corresponding metadata records. Easy browsing to the wide range of the available content was achieved using a tag cloud feature that was developed from scratch for the purposed of the project. The tag cloud was applied to the thematic category data field, which used the EuroVoc thesaurus as a controlled vocabulary for assigning thematic categories to individual items. Regarding digital material presentation, a streaming video player was incorporated into the item pages to enable easy access to the video content and including configuration capabilities that enabled differentiation of the player presentation based on values of specific metadata fields. Regarding text documents, an online reading feature was included in the repository, enabling users to open, read online and search into big texts without the need to save them in their systems – an implementation fully compatible with tablets like iPad and Android-based devices. The infrastructure that enables that is based on transforming documents to JPEG2000 images, including OCR text for searching, and serving them to users via an extended version of the Internet Archive Book reader using open source and W3C standards compatible technologies.

### Bionotes

**Ioanna-Ourania Stathopoulou** received a B.Sc. in Computer Science and, later on, a Ph.D. degree from the University of Piraeus, Piraeus, Greece. Her doctoral research was sponsored by the General Secretary of Research and Technology of the Greek Ministry of Development, under the auspices of the PENED-2003 basic research program. Since May 2007, she is with the Department of Software Application Development of the National Documentation Centre (EKT) of the Hellenic Research Foundation, where she works as a software engineer participating in the design and implementation of digital repositories, e-publishing platforms and bibliographic systems. Her primary research interests are in the areas of affective computing, human-computer interaction, computer vision, and pattern recognition, and their applications in user modeling, information retrieval and intelligent software systems. She has over 30 publications in peer reviewed journals and conferences and has co-authored a monograph entitled “Visual Affect Recognition”, published by IOS Press. Email: [iostath@ekt.gr](mailto:iostath@ekt.gr)

**Nikos Houssos** works at the National Documentation Centre/NHRF in Athens, Greece as Head of the Software Development Unit. He is the software architect of the Greek “National Information System on Research and Technology” (EPSET) which comprises a variety of scholarly communications systems such as CRIS, repositories, e-publishing platforms and bibliographic systems. He has designed a number of grey literature systems including the Hellenic National Archive of Doctoral Dissertations. He participates in various EU FP7 projects like OpenAIRE/OpenAIREPlus, Arrow Plus, ENGAGE and PAERIP. He is a member of the euroCRIS Board and a contributor to the development of the CERIF data model. He has participated in EU FP5 and FP6 IST projects related to mobile networking and services (1999-2004), and lectured at the Technical University of Crete (2004-2007). He holds a

## Bionotes (continued)

Ph.D. in Computer Science from the University of Athens and has co-authored more than 30 peer-reviewed publications in international journals and conferences. Email: [nhoussos@ekt.gr](mailto:nhoussos@ekt.gr)

**Dr. Panagiotis Stathopoulos** received his diploma in Electrical and Computer Engineering and his PhD in Broadband Networks at 1999 and 2004 respectively, from the National Technical University of Athens (NTUA). From 1999 until 2006 he has been with the Computer Networks Laboratory of NTUA participating and technically coordinating research projects in the areas of broadband communications and applications. From 2006, he is the head of the Systems and Networks Unit of EKT leading the team developing a highly sophisticated IT infrastructure, for providing advanced open access applications and services. He has taught at the University of the Aegean and the TEI of Piraeus, and he has over 30 publications in peer reviewed journals and conferences. Email: [pstath@ekt.gr](mailto:pstath@ekt.gr)

**Despina Hardouveli** is a senior Information Professional in Greece having an experience of nearly 30 years in the field. She graduated in Biology and holds a master degree. Since 1983 she works for National Documentation Centre (NDC/NHRF) in Greece. From 1993 to 2006 was Head of the S+T Information Services Department of NDC/NHRF, responsible for the provision of scientific & technological information to the Greek research and academic community. Since 1997, she has been a member of the project group for National Information System for Research and Technology (NIRST), a project elaborated within the NSRF (National Strategic Reference Framework) and the European Operational Program, responsible for several tasks, among them the settlement of international databases into the NDC's information systems, the support to the creation of the Electronic Reading Room, the transition from traditional retrieval services to digital content services, the conceptual design and content management of special bibliographic collections, the digital library integration. In charge of the creation of the NHRF's OA Repository "Helios", she is responsible for the content maintenance and upgrading, providing supportive services to the stakeholders/users. She was instrumental in the development of a funder repository of mainly grey literature material of diverse types, produced under the funding programmes of the Hellenic Ministry of Education (co-financed by the European Union). She has served on numerous project committees and was demonstrator & trainer of the digital content services of NDC. From 2004 to 2010 was a member of GRNET S.A (Greek Research & Technology Network) Administrative Board. Email: [dxardo@ekt.gr](mailto:dxardo@ekt.gr)

**Alexandra Roubani** holds a Bachelor of Science (BSc), Technological Educational Institute of Athens, Department of Librarianship and Information Systems, with a Master's Degree in Library Sciences (MLIS), Ionian University, Department of Archives and Library Science. She also studied Cultural Administration in the Department of Communication, Media and Culture, Panteion University of Social and Political Sciences (BSc). She is a cataloguing and metadata expert at Institutional Repositories of the National Documentation Centre (EKT)/National Hellenic Research Foundation since 1997, and her main occupation is focused on the gradual alteration/transformation of the National Union Catalogue of Serials. Her professional interests also include Authority Files, Digital Preservation Metadata Standards, Interlibrary Loan and Digital Curation. Email: [arouba@ekt.gr](mailto:arouba@ekt.gr)

**Ioanna Sarantopoulou** is the Head of the Digital Library Department of the National Documentation Centre (EKT) in the National Hellenic Research Foundation (NHRF). She is involved in organizing and maintaining Greek digital content in Library's collection development and also in providing intermediated library information services. She participated in several workgroups within the EKT for the implementation of the National Information System for Research and Technology. She formerly worked for many years at the EKT, using online systems for documentation and information retrieval, to facilitate information seeking for scientists on Natural Sciences and Engineering. She holds a Bachelor's degree from National Technical University of Athens in Chemical Engineering. Email: [isaran@ekt.gr](mailto:isaran@ekt.gr)

**Alexandros Soumplis** studied "Engineering in Informational and Communication Systems" and holds a "MSc in Network & Data Communication" from Kingston University UK. Since 2010 is accepted by the Faculty of Sciences of the Hellenic Open University as a PhD student. His research interests involve informal learning environments as well as the use of innovative technologies for learning. Furthermore he has more than 12 years of active experience as a systems engineer with focus on core IT systems and in the past has worked for major computer and telecommunications companies in Greece. Since 2007 collaborates with the National Documentation Centre as a member of the "Information Systems and Networks Department" and has active involvement in several projects. Also, through his academic and professional career has submitted work and participated in several scientific and business conferences. Email: [soumplis@ekt.gr](mailto:soumplis@ekt.gr)

**Chrysostomos Nanakos** has extensive experience in Open Source technologies. He has successfully executed several projects with different complexities and sizes in the Public and Private sector based on Open Source architectures. Since 2011 he cooperates with the National Documentation Center as a Senior Systems Administrator and Open Source Developer. He received his diploma in Electrical and Computer Engineering from the National Technical University of Athens (NTUA) where he is currently working towards the Ph.D. degree in Computational and Applied Electromagnetics. Email: [cnanakos@ekt.gr](mailto:cnanakos@ekt.gr)

## A funder repository of heterogeneous grey literature material with advanced user interface and presentation features

Ioanna-Ourania Stathopoulou, Nikos Houssos, **Panagiotis Stathopoulos**,  
Despina Hardouveli, Alexandra Roubani, Ioanna Sarantopoulou, Alexandros  
Soumplis, Chrysostomos Nanakos;

National Documentation Centre / National Hellenic Research Foundation, Greece



GL14 – "Tracking Innovation through Grey Literature", Rome, 29-30 November 2012

## The National Documentation Centre (EKT)

- The national organisation in Greece:
  - for scientific documentation, online information and support services on research, science and technology
  - Objective: **making knowledge accessible to everyone**
- Incorporated in National Hellenic Research Foundation (NHRF)
- Implements the "National Information System for Research and Technology": <http://epset.gr>
  - CRIS systems, Digital Repositories, e-Publishing, Digital Libraries, Interactive Culture, and more...
  - Open Access advocate: <http://openaccess.gr>, OPENAIRE, OPENAIRE+ member and NOAD.
  - Greek PhD theses archive: <http://phdtheses.ekt.gr>

GL14 – "Tracking Innovation through Grey Literature", Rome, 29-30 November 2012



## Project context

- Hellenic Ministry of Education programmes for education and lifelong learning, co-funded by the EU structural funds
  - Three phases: 1994-1999, 2000-2006, 2007-2013
  - Overall funds in the range of billion Euros
- Activities funded:
  - Research: theses, collaborative projects
  - Studies, surveys, reports, project public deliverables
  - Conferences, workshops
  - Face-to-face courses / seminars
  - Creation of educational material
- Most of them fall into the realm of "grey literature"

GL14 – "Tracking Innovation through Grey Literature", Rome, 29-30 November 2012



## Aims and scope of the project

- Decision by funding authority to create a repository (autumn 2010)
- Aims: Dissemination, reuse and preservation of the digital material produced in the frame of the funded programmes
  - Single point of (persistent) access to material
  - Standards-based documentation/metadata, interoperability
  - Link of content to project / funding programme (by an external CRIS system)
- Repository available (since spring 2011)
  - <http://repository.edulll.gr>

GL14 – "Tracking Innovation through Grey Literature", Rome, 29-30 November 2012




---

---

---

---

---

---

---

---

## Project inputs and outputs

- Input material:
  - Digital files arranged in folders – (hopefully) according to project
  - List of projects
  - No metadata!
- Output aimed for:
  - Repository with standards compliant metadata records along with processed digital files, suitable for dissemination
  - Internal lightweight CRIS system combining digital items with project information and metadata
  - User-friendly interface for human users
  - Programmatic interface (OAI-PMH) for reuse by third-party software applications

GL14 – "Tracking Innovation through Grey Literature", Rome, 29-30 November 2012




---

---

---

---

---

---

---

---

## Methodology and approach

- Investigate input material
  - Identify content of archival value – exclude all other material (e.g. administrative documents)
  - Assign priorities to content
- Documentation / cataloguing of material according to standards – application profile created
  - Subject cataloguing
  - No self-archiving at this phase
- Digital processing of input files to produce copies suitable for preservation and dissemination – both initial and processed files kept in repository
- A lot of effort allocated to user-friendly interfaces for browsing metadata and accessing content

GL14 – "Tracking Innovation through Grey Literature", Rome, 29-30 November 2012




---

---

---

---

---

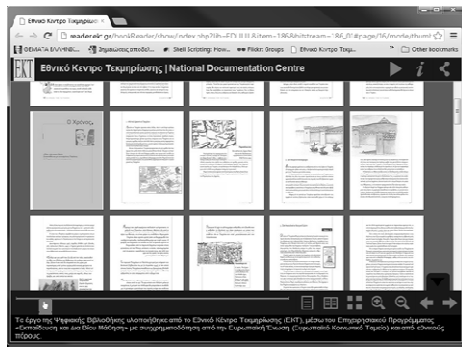
---

---

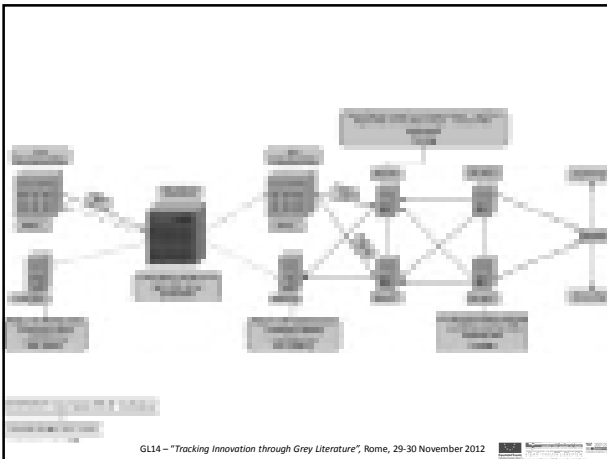
---



Online reader (screenshot)



GL14 – “Tracking Innovation through Grey Literature”, Rome, 29-30 November 2013



## Future work

- Support self-archiving (two phase cataloguing)
- Integrate CERIF/CRIS functionality to the publicly available main system
- Various new features for presentation of content

GL14 – “Tracking Innovation through Grey Literature”, Rome, 29-30 November 2012



## Working for an open e-publishing service to improve grey literature editorial quality

**Rosa Di Cesare and Marianna Nobile**, Institute for Research on Population and Social Policies, IRPPS  
**Silvia Giannini**, Institute of Information Science and Technology, ISTI, Italy

The widespread diffusion of electronic publishing technologies is creating a “second life” for traditionally grey documents as well as an innovative way of managing grey contents. Academic and research libraries that have had a fundamental role in supporting Open access practices in the construction of Institutional repositories and digitization programs are currently moving toward the development of additional services for their community scholars. This is in line with the necessity of libraries reshaping their role in the digital age following changes in the scholarly communication models. In this context library publishing services represent a new modality to diffuse scholarly research outputs, improve the quality of in-house published products and decrease costs of publication.

At IRPPS an e-publishing service based on Open Journal Systems (OJS) was developed to manage GL collections. To its development an analysis of IRPPS current practices of publishing was carried out considering types of documents and contents to be selected for future e-publications as well as monographs and digitization of previously published or unpublished works that have represented important achievements of IRPPS research results. A new editorial plan was performed in collaboration with the internal scientific community to define roles of the changed publishing process as well as editorial policies aimed to improve the scientific quality and visibility of IRPPS research products.

Starting from this experience, the paper describes a survey on the editorial practices carried out at CNR Research Institutes in order to verify whether IRPPS pilot project can become a starting point for the development of e-publishing services in other CNR Institutes and/or Departments.

The survey analysis the GL production available at CNR Institute Web sites and focuses in particular on the identification of:

- Publishing profile (publication frequency and type of editorial production);
- Level of dissemination (i.e. in international archives and/or national catalogues);
- Editorial policies and practices (scientific and editorial board, open access policy, peer-review system and/or other reviewing system);
- Use of bibliographic elements and International bibliographic standards.

### Bionotes

**Rosa Di Cesare** is responsible for the library at the Institute for research on populations and social policies of the National Research Council (CNR). She worked previously at the Central library of CNR where she became involved in research activities in the field of Grey literature (GL) as member of the Technical Committee for the SIGLE database. Her studies have focused on the use of GL in scientific publications and recently on the emerging models of scholarly communication (OA and IR). Email: [r.dicesare@irpps.cnr.it](mailto:r.dicesare@irpps.cnr.it)

**Marianna Nobile** graduated in Linguistics, Faculty of Letters and Philosophy at the University of Rome La Sapienza. She had an internship in the library of the Senate of the Republic (Giovanni Spadolini Library), where she worked with the Books Acquisition. Currently she is collaborating with the library of the Institute for Research on Population and Socials Science of the Italian National Research Council (CNR) and she is involved in the activities of indexing electronic resources as well as in the development of an e-publishing service. Email: [marianna.nobile@irpps.cnr.it](mailto:marianna.nobile@irpps.cnr.it)

**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](mailto:silvia.giannini@isti.cnr.it)

## Working for an open e-publishing service to improve grey literature editorial quality

**Rosa Di Cesare, Marianna Nobile**

*Institute for Research on Population and social policies, IRPPS, Italy*

**Silvia Giannini**

*Institute of Information Science and Technology, ISTI, Italy*



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council  
Rome, Italy, 29-30 November 2012



### Background of the study

#### Advantages in the development of e-Publishing service

- ◊ Sustainable economic model based on Open source technologies
- ◊ Enhancement of the editorial quality of "in-house" publications
- ◊ "Second life" for previously paper-based Grey documents
- ◊ New opportunities to manage and diffuse research outputs
- ◊ Providing services that researchers understand, need and value
- ◊ Suitable for Humanities and Social Science (locally-oriented and mainly monograph-based)



**New role of Academic and Research Libraries in the co-management and integration of different services?**  
(Repositories, Digitization initiatives, Digital preservation)



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council  
Rome, Italy, 29-30 November 2012



### Aims

#### Analysis of CNR Institutes' editorial production

- ❖ type of products
- ❖ editorial quality
- ❖ technology used
- ❖ type of distribution
- ❖ access and discovery



**Improve editorial quality of in-house scientific publications**



**Increase availability of scholarly research outputs at a reasonable costs**



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council  
Rome, Italy, 29-30 November 2012



## Methods

**Source of analysis:** CNR Research Institutes website

**Data analysed:** products published in-house and/or in collaboration with commercial publishers

### Type of analysis: qualitative

- Identification of current and ceased products with a minimal set of editorial elements (i.e.: series title and/or number);
- Classification of products according to their level of innovation in content management
- Identification of the business model
- Identification of a set of indicators to measure products' editorial quality

**Period of analysis:** data was gathered in September 2012



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council, CNR  
Rome, Italy, 29-30 November 2012



## Survey design

Departments	No. of institutes	I	A	Z	X
Earth and environment	13	6	8	7	0
Agriculture and food sciences	10	2	6	7	2
Biomedical sciences	17	0	1	3	13
Chemistry and material techn. sciences	14	0	0	7	7
Physics sciences	14	0	2	5	7
Engineering and ICT	21	2	9	10	2
Social sciences and humanities	20	9	6	5	11
<b>TOTAL</b>	<b>109</b>	<b>19</b>	<b>87</b>	<b>44</b>	<b>42</b>

### Criteria for the selection of the Institutes:

I = CNR Institutes that manage its in-house scientific production in an innovative way  
A = CNR Institutes that manage its in-house production applying an editorial control that includes at least a standardized series title and number

← Analyzed Institutes

Z = CNR Institutes that produce GL without applying any editorial control  
X = CNR institutes that have no GL production

← Excluded Institutes



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council, CNR  
Rome, Italy, 29-30 November 2012



## The data

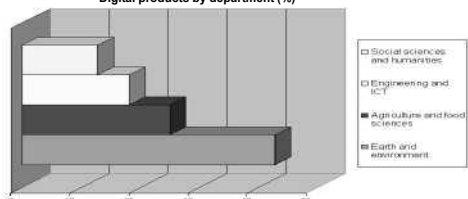
❖ 106 products analyzed:

- ✓ 87 traditional products
- ✓ 19 digital products

### Traditional and digital editorial products by Department

Departments	No. of institutes	Total number of traditional products	Number of digital products
Earth and environment	13	14	8
Agriculture and food sciences	10	8	2
Biomedical sciences	17	1	1
Chemistry and material techn. sciences	14	3	3
Physics sciences	14	2	3
Engineering and ICT	21	11	2
Social sciences and humanities	20	10	9
<b>TOTAL</b>	<b>109</b>	<b>87</b>	<b>19</b>

Digital products by department (%)



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council, CNR  
Rome, Italy, 29-30 November 2012



### Analysis of the business process

**NO Linear process:**  
 Depending on the business model adopted  
 on the organisational framework  
 on the type of products  
 on the access policies  
 on the evaluation strategies  
 on the technology used  
 .....

.... Some examples

GL14 - Fourteenth International Conference on Grey Literature  
 National Research Council, CNR  
 Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

### From traditional to digital products 1)

**Business model:**  
 In-house production & distribution  
 Peer-review: internal evaluation for WPs, external for Monographs  
 Full OA  
 OJS platform

GL14 - Fourteenth International Conference on Grey Literature  
 National Research Council, CNR  
 Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

### From traditional to digital products 2)

*From a high quality GL to an OA journal*

**Business model:**  
 In-house production  
 Peer-reviewed Journal  
 External copy-editing  
 Full OA from 1999  
 OJS platform

GL14 - Fourteenth International Conference on Grey Literature  
 National Research Council, CNR  
 Rome, Italy, 29-30 November 2012

---

---

---

---

---

---

---

---

### From traditional to digital products 3)



*From paper-based to electronic journal*

**Business model:**  
International disciplinary community  
Peer-reviewed Journal  
International commercial publisher



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council  
Rome, Italy, 29-30 November 2012




---

---

---

---

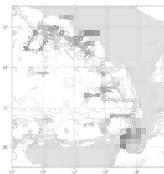
---

---

---

---

### From traditional to digital products 4)



**Business model:**  
In-house production  
Full OA



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council  
Rome, Italy, 29-30 November 2012




---

---

---

---

---

---

---

---

### Conclusions

#### Editorial products

- ❖ In-house products vary according to disciplinary fields
- ❖ In Science and Humanities the highest number of traditional editorial products
- ❖ Digital products are concentrated in Earth and Environmental Sciences

#### Business models

- ❖ Disciplinary fields do not influence the business model
- ❖ The adoption of new technologies does not depend on the disciplinary field, but is influenced by a long tradition in editorial production

#### Editorial quality

- ❖ Stability, visibility, adoption of bibliographic and editorial standards increase the quality of both traditional and digital products



GL14 - Fourteenth International Conference on Grey Literature  
National Research Council  
Rome, Italy, 29-30 November 2012




---

---

---

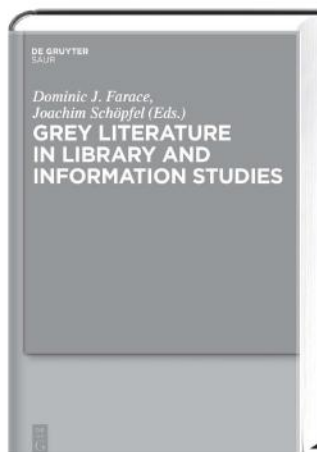
---

---

---

---

---



*Dominic J. Farace and Joachim Schöpfel (Eds.)*

## **GREY LITERATURE IN LIBRARY AND INFORMATION STUDIES**

2010. vi, 282 pages

**Hardcover** RRP € 89.95 [D] / US\$ 126.00. ISBN 978-3-598-11793-0

**eBook** RRP € 89.95 / US\$ 126.00. ISBN 978-3-598-44149-3

The further rise of electronic publishing has come to change the scale and diversity of grey literature facing librarians and other information practitioners. This compiled work brings together research and authorship over the past decade dealing with both the supply and demand sides of grey literature. While this book is written with students and instructors of Colleges and Schools of Library and Information Science in mind, it likewise serves as a reader for information professionals working in any and all like knowledge-based communities.

### **CONTENTS**

Introduction Grey Literature (*Farace and Schöpfel*)

#### **Part I – Producing, Processing, and Distributing Grey Literature**

##### **Section One: Producing and Publishing Grey Literature**

**Chapter 1** Grey Publishing and the Information Market: A New Look at Value Chains and Business Models (*Roosendaal*)

**Chapter 2** How to assure the Quality of Grey Literature: The Case of Evaluation Reports (*Weber*)

**Chapter 3** Grey Literature produced and published by Universities: A Case for ETDs (*Južnič*)

##### **Section Two: Collecting and Processing Grey Literature**

**Chapter 4** Collection building with special Regards to Report Literature (*Newbold and Grimshaw*)

**Chapter 5** Institutional Grey Literature in the University Environment (*Siegel*)

**Chapter 6** Copyright Concerns Confronting Grey Literature (*Lipinski*)

##### **Section Three: Channels for Access and Distribution of Grey Literature**

**Chapter 7** Theses and Dissertations (*Stock and Paillassard*)

**Chapter 8** Grey Documents in Open Archives (*Luži*)

**Chapter 9** OpenSIGLE – Crossroads for Libraries, Research and Educational Institutions in the Field of Grey Literature (*Farace, Frantzen, Stock, Henrot, and Schöpfel*)

#### **Part II – Uses, Applications, and Trends in Grey Literature**

##### **Section Four: Applications and Uses of Grey Literature**

**Chapter 10** The driving and evolving Role of Grey Literature in High-Energy Physics (*Gentil-Beccot*)

**Chapter 11** The Use and Influence of Information Produced as Grey Literature by International, Intergovernmental Marine Organizations: Overview of Current Research (*MacDonald, Wells, Cordes, Hutton, Cossarini, and Soomai*)

**Chapter 12** Grey Literature in Karst Research: The Evolution of the Karst Information Portal, KIP (*Chavez*)

**Chapter 13** Grey Literature Repositories: Tools for NGOs Involved in Public Health Activities in Developing Countries (*Crowe, Hodge, and Redmon*)

##### **Section Five: Future Trends in Grey Literature**

**Chapter 14** Blog Posts and Tweets: The Next Frontier for Grey Literature (*Banks*)

**Chapter 15** Assessing the Return on Investments in Grey Literature for Institutional Repositories (*Schöpfel and Boukacem*)

**Chapter 16** e-Science, Cyberinfrastructure and CRIS (*Jeffery and Asserson*)

**Chapter 17** Course and Learning Objective in the Teaching of Grey Literature: The Role of Library and Information Science Education (*Rabina*)

# **Teaching and learning across national frontiers to improve dissemination of scientific research output: Is there a borderline between grey and white literature? Lessons learned from the NECOBELAC experience in Europe and Latin America**

**Paola De Castro**, Istituto Superiore di Sanità, Italy

Now that both grey and white literature freely circulates on the web, the traditional borderline between the two genres becomes more and more indistinct. Some of the characteristics traditionally associated with GL in the pre-Internet age (such as limited circulation, poor editorial quality, absence of peer review), do not apply to the GL of the 21<sup>st</sup> century. Yet, also white literature (mainly scientific journals and e-books) is rapidly changing and becomes closer to GL in its new and often controversial vocation of being open and free on the Internet even in its earliest stages, according to different OA publishing models (such as ahead-of-publication articles or new OA experimental journals, “un-journals”, or e-books).

A new form of scientific conversation including both traditionally white or grey literature is going to replace old categories; according to the most recent trends in information policies at European and international level, authors are also asked to link scientific data (data sets) to their research articles, thus accelerating the progress of science by further sharing existing information and adding grey nuances to white literature. Moreover, alternative dissemination patterns (namely through mobile technologies and social networks) are rapidly being developed and widely utilized, initially only by the youngest generations and now by all age groups at global level.

In this context, the production and dissemination of scientific contents on the web requires new levels responsibilities and skills for all stakeholders (primarily authors of scientific publications, editors and librarians). In most cases, training is necessary to create awareness on the different ethical and technical implications associated with open access publishing, including both white and grey literature.

The training experience carried out in this regard within the NECOBELAC project ([www.necobelac.eu](http://www.necobelac.eu)) will be reported with special focus on the difference between white and grey literature (if any).

NECOBELAC (NEtwork of COllaboration between Europe and Latin America and Caribbean countries) is a project funded by the European Commission within the 7th Framework program (Science in Society) in the period February 2009 - July 2012. The project partners belong to academic and scientific institutions in Europe and Latin America (Istituto Superiore di Sanità, Italy, project coordinator; The University of Nottingham, United Kingdom; Consejo Superior de Investigaciones Científicas, Spain; BIREME, PAHO/WHO, Brazil; Instituto de Salud Pública, Colombia; Universidade do Minho, Portugal).

The main objectives of the NECOBELAC project were to: improve scientific writing; create awareness of open access publishing, and establish durable collaborations among academic and scientific institutions from European and Latin American countries in the field of public health. These objectives were achieved through three main actions including the development and implementation of:

- a two-level training strategy including courses for trainers (T1) and local training replication activities (T2);
- innovative training tools (NECOBELAC topic maps) to be utilized in the training activities above;
- a network of health institutions and ICT supporting systems in Europe and Latin America to support the training activity, promote health information dissemination and stimulate new research collaboration on specific public health topics.

The results achieved by the NECOBELAC project will be presented with special reference to GL implications.

In brief, NECOBELAC carried out 46 training initiatives (T1 and T2), created a network of over 200 supporting institutions in Europe and Latin America developed original topic maps on scientific writing and open access publishing.

NECOBELAC topic maps based on semantic web technologies represent an innovative training support tool. GL is included as a specific topic in the NECOBELAC maps, and most maps and associated schemes are equally applicable to both open and grey literature. Examples of the topic maps will be shown and as well as lessons learned from the NECOBELAC experience in Europe and Latin America.

## Bionote

**Paola De Castro**, Head of the Publishing Unit of the Istituto Superiore di Sanità ([www.iss.it](http://www.iss.it), the National Institute of Health in Italy) and member the editorial committees of the publications edited by the Institute (Annali dell'Istituto Superiore di Sanità, a quarterly peer-reviewed science journal, Notiziario dell'Istituto Superiore di Sanità, a monthly newsletter on current research and other activities, different series of technical reports and of the Journal European Science Editing and Eurosurveillance. She fostered the adoption of an institutional policy for Open Access and now the Digital archive of the ISS (<http://dspace.iss.it/dspace/>), including the publications of ISS researchers and other health institutions in Italy, has become the most important digital archive in health sciences in Italy.

Coordinator of the European FP 7 Project NECOBELAC ([www.necobelac.eu](http://www.necobelac.eu)), a network of European and Latin American institutions involved in training activities in scientific writing and open access publishing.

Published many articles on information transfer process and research evaluation; delivers courses on scientific writing and new publication models at national and at international level. She organizes exhibitions on public health issues addressed to both specialized staff and to the general public, including students. She is involved in a number of activities for the promotion of new publications models and science culture in schools; she also carries on projects regarding the preservation of historical heritage belonging to scientific institutions (including scientific instruments no longer used for research purposes, unpublished documents, pictures, oral memories). Email: [paola.decastro@iss.it](mailto:paola.decastro@iss.it)



Teaching and learning across national frontiers  
to improve dissemination of scientific research output

1

## IS THERE A BORDERLINE BETWEEN GREY AND WHITE LITERATURE?

Lessons learned from the NECOBELAC  
experience in Europe and Latin America

**Paola De Castro**  
Istituto Superiore di Sanità, Rome (Italy)  
., 2012

14<sup>th</sup> International Conference on GL • Tracking Innovation through Grey Literature  
Rome, November 29-30, 2012

---

---

---

---

---

---

---

---

## Grey and white literature today

How many other coulours can be used?

2

<b>GL before the Internet</b> Limited circulation	<b>Today</b> GL and white literature (WL) freely circulate on the web
<b>Poor editorial quality</b>	<b>The borderline between them</b> becomes more and more unclear
<b>Difficult to retrieve</b>	<b>WL is closer to GL and viceversa</b> in their new and often controversial
<b>Addressed to small circles</b>	<b>vocation of being open and free</b> according to different OA publishing models
<b>Difficult to define</b>	

14th International Conference on Grey Literature Rome November 29-30, 2012

---

---

---

---

---

---

---

---

## New forms of scientific conversation

3

**HowOpenStk?**

**Old categories are being replaced**

- authors are asked to deposit their papers in repositories and **link scientific data** (data sets) to their research articles
- **alternative dissemination patterns** (mobile technologies and social networks) are developed and utilized by all age groups at global level.

14th International Conference on Grey Literature Rome November 29-30, 2012

---

---

---

---

---

---

---

---

## New levels of responsibility arise for all stakeholders

4

### Training is necessary

- to **create awareness** on the ethical and technical implications associated with OA publishing, including both white and grey literature.
- to **provide tools** to cope with the new requirements associated with such new responsibilities



NECOBELAC project

14<sup>th</sup> International Conference on Grey Literature • Rome November 29-30, 2012

---

---

---

---

---

---

---

---

## NECOBELAC project in brief

Network of Collaboration Between Europe & Latin American-Caribbean countries

It is a European project (2009-2012)  
"Science in Society" working in the field  
of PUBLIC HEALTH



www.necobelac.eu

It aims to:

- ✓ improve scientific writing
- ✓ promote open access publishing models
- ✓ foster technical and scientific cooperation between Europe and Latin America



All actors taking part in  
scientific communication process

14<sup>th</sup> International Conference on Grey Literature • Rome November 29-30, 2012

---

---

---

---

---

---

---

---

## NECOBELAC and GL

6



www.necobelac.eu

- The **training experience** carried out within the **NECOBELAC project (2009-2012)** will be reported with special focus on the difference between white and grey literature.
- NECOBELAC methodology can be usefully applied to GL community

14<sup>th</sup> International Conference on Grey Literature • Rome, November 29-30, 2012

---

---

---

---


---

---


---

---

## NECOBELAC key points



- MAIN ACTIONS**
  1. Two level training activity *interrelated*
  2. Innovative training tools
  3. Networking & cooperation activity
- BI-DIRECTIONAL APPROACH**  
sharing experiences and initiatives between Europe & Latin America within a common strategy
  - This approach offers mutual benefits EU-LAC and supports cultural change in production, use and dissemination of health information
- All project partners contributed to the three main actions (participating in training, provide contents for training tools, support the network)




---

---

---

---

---

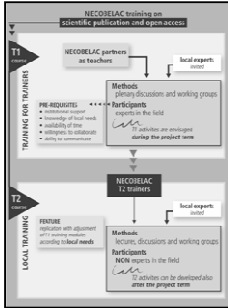
---

---

---

## Action 1

### TWO-LEVEL TRAINING ACTIVITY



**T1**

**TRAINING COURSES FOR TRAINERS**

8 courses in EU and LAC

**T2**

**TRAINING REPLICATION**

local training courses, workshops, meetings, etc.  
40 replications to date

14<sup>th</sup> International Conference on Grey Literature • Rome November 29-30, 2012

---

---

---

---

---

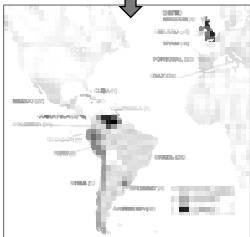
---

---

---

## Figures from NECOBELAC training

**8 Training for trainers (T1)**



Brazil	Sao Paulo	2010
Italy	Rome	2010
Colombia	Bogotá	2010
Spain	Madrid	2011
Argentina	Buenos Aires	2011
Portugal	Braga	2011
Mexico	Guadalajara	2011
Ireland	Dublin	2012

**40 Local training replications (T2)**

**15 countries in the network**

**4 Europe + 11 Latin America**

**>200 institutions**

**Lessons learnt**  
Be flexible, provide support and teaching material

---

---

---

---

---

---

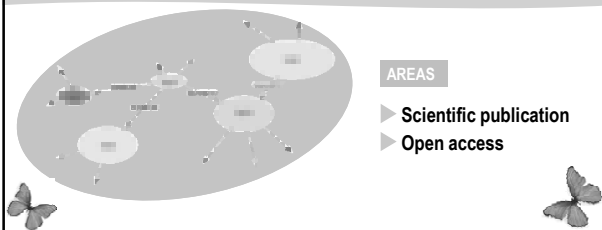
---

---

## Action 2

### INNOVATIVE TRAINING TOOLS NECOBELAC topic maps

Topic maps are based on semantic web technology  
They associate topics and link them to selected online resources  
They have a triple SUBJECT-PREDICATE-OBJECT structure




---

---

---

---

---

---

---

---

### NECOBELAC topic maps

#### Scientific publication




---

---

---

---

---

---

---

---

### NECOBELAC topic maps

#### Open access




---

---

---

---

---

---

---

---

## Grey literature in Australian education

**Gerald White, Julian Thomas, Paul Weldon, Amanda Lawrence and Helen Galatis**  
Australian Council for Educational Research, Australia

The prevalence of informal publishing or grey literature in education would appear to have increased as digital technologies have become mainstream, educators have become more proficient and policies have moved increasingly towards supporting its use. In addition, the take up of social networking technologies and innovative methods of digital publishing have encouraged educators to produce, distribute and share content and commentary. Grey literature may make a substantial contribution to education even though issues such as credibility, access and a lack of standards can pose problems for producers and users. This paper will begin by providing a context for the discussion within the broader policy and education environment in Australia and be followed by an overview of grey literature as it appears in education in Australia introducing evidence of its usage, dissemination and application in Australian education. Evidence about the access, dissemination and use of grey literature will be drawn from an examination of the characteristics of a number of dominant social networking and digital publishing services that are used by educators in schools, training institutes and teacher education faculties. This evidence will be discussed in the context of influential national, state and institutional policies that address the use of digital technologies in education. As the take up of digital technologies in education increases, there is an expectation that the access to, dissemination of and use of digital publishing by educators and for educators will increase and have an impact on online professional learning and awareness of education research and practices.

### Bionotes

**Dr. Gerald (Gerry) White** is a Principal Research Fellow at the Australian Council for Educational Research (ACER). He specialises in the use of digital technologies and digital media in education and currently manages the Digital Education Research Network (DERN) (<http://www.dern.org>) which publishes weekly research reviews about ICT in education. Formerly head of Australia's education technology national agency for education and training, Gerry's interests and experience are in digital diffusion, online collaboration, grey literature, teaching and learning, leadership and online communities. Email: [whiteg@acer.edu.au](mailto:whiteg@acer.edu.au)

**Professor Dr. Julian Thomas** has been the Director, Institute for Social Research (ISR), Swinburne University of Technology since 2005. In that role he has been responsible for a large number of publications about the influence of the Internet, government policies and social issues as well as part of the World Internet Project. Professor Thomas has initiated and led a number of online information and commentary services such as Australia Policy Online. Professor Thomas is one of the Chief Investigators of a major national Australian research project into Grey literature, policy innovation and access to knowledge about realising the value of informal publishing.


**Dr. Paul Weldon** completed his PhD in sociolinguistics in 2007 and joined ACER in January 2010. For the three years prior to this he was Research Associate at Independent Schools Victoria. Dr Weldon has had a diverse career in Australia, the UK and in China, where he worked for two years as English Language Editor for the Journal of China University of Geosciences. He has worked on research projects in the fields of Social Capital, Education for Sustainability, School-Community Partnerships, Cybersafety, and on the development of school performance measures. He has published several reports and articles and he has also been a member of the Australian Cybersafety Consultative Group.

**Ms. 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](mailto:alawrence@swin.edu.au)

**Ms. Helen Galatis** is a Research Fellow with the Australian Council for Educational Research (ACER) responsible for research, information retrieval, online aggregation and publishing for the Digital Education Research Network. Formerly, Ms Galatis has managed all aspects of the online component of the Education Network Australia service for the Australian Government. Ms Galatis specialises in online communities, web services, metadata and archiving.

**ACER**

## Grey literature in Australian education



*Dr Gerald White  
(Gerry)*

Principal Research Fellow  
Digital Education Research Network

<http://dern2.acer.edu.au>

---

---

---

---

---

---

---

---

**ACER**

## ARC research grant

**Grey Literature, policy innovation and access to knowledge:  
realising the value of informal publishing**

<p><b>Universities</b> Swinburne University (Administration) - Institute for Social Research Victoria University</p> <p><b>Partner Organisations</b></p> <ul style="list-style-type: none"> <li>Eidos Institute</li> <li>Australian Council for Educational Research</li> <li>National Library Australia</li> <li>National and State Libraries Australasia</li> </ul>	<p><b>Chief Investigators</b> Professor Julian Thomas Professor John Houghton</p> <p><b>Key Informant</b> Dr Gerald White</p> <p><b>Research Associates</b> Amanda Lawrence Dr Paul Weldon</p>
---	--

---

---

---

---

---

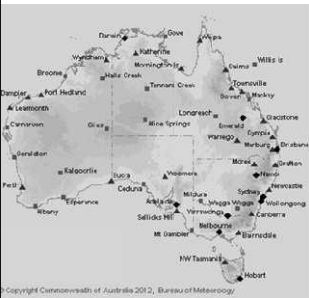
---

---

---

**ACER**

## Overview



- Grey literature
- My EdNA story
- Things changed
- Information abundance
- Lost information
- Integrators/curators
- The Tyndall framework

---

---

---


---

---

---

---

---




## EdNA's life

1. Antecedent OLTC (1993-1997)
2. EdNA development commenced 1995
3. Life 1997-2011 (14 years)

AICTEC

**Collaboration**

- 8 State departments schools (65%)
- Catholic education (23%)
- Independent education (12%)
- 8 Training departments
- 38 Universities
- Commonwealth education
- Agencies – technology & resources



---

---

---


---

---

---

---

---



## EdNA usage

**Resource database** (*search and browse*)

Total database items = 41,368, ( **39,694** resources, 871 events, 803 news items)

Resource items accessible by distributed search = excess of **3 million** digital resources

**Memberships** (*Shared Information Services*)

35,349 members have self-registered with any edna registered service via single sign on (i.e. Groups, Lists administration, me.edu.au or any combination)

23,205 edna Groups (Moodle) members

14,729 me.edu.au members

123,928 edna List email subscribers

**Groups** (*EdNA Groups*)

Groups *member* growth over 2 year period Jan 2007-December 2008 = **113%**

Growth in number of *communities* for same period = **100%**

Growth in *single sign-on* (i.e. all member services) for same period = **127%**

Growth in email *subscribers* for same period = **65%**.

(Education.au, 2009b, pp. 19-20)

---

---

---


---

---


---

---

---

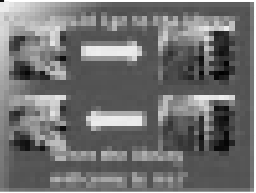


## Things changed



96% 9-16 year olds access the internet often using mobile devices (60%)

ACMA. (2011). *Communications report 2010-11 series: Report 3 – The emerging mobile telecommunications service market in Australia*. ACMA: Canberra, Australia. Retrieved December 20, 2011, from [http://www.acma.gov.au/WEB/STANDARD/pc=PC\\_410225](http://www.acma.gov.au/WEB/STANDARD/pc=PC_410225).



---

---

---


---

---


---

---

---




## Education and abundance




**3.5 million Australians**  
**Annual budget ~ \$80 bn**

### Abundance – blogs informally published

Powering **1,446,974** Edublogs since 2005





Read as: In October 2011 NM Incite tracked 973 million blogs as sources of online heat.  
Source: NM Incite

NM Incite. (2012). Number of blogs tracked by NM Incite. Retrieved march 14, 2012, from [http://www.nmincrite.com/?page\\_id=210](http://www.nmincrite.com/?page_id=210)

---

---

---

---

---

---

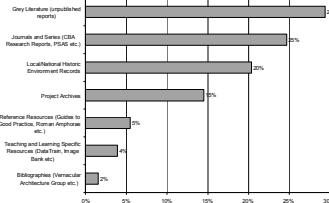
---

---

## Abundance of information

### Archaeological Research Data

(Per cent of users by last data type used)



Data Type	Percentage
Grey Literature (unpublished reports)	30%
Journals and Series (CSA, Research Reports, PLOS, etc.)	25%
Local/Regional/National Environment Records	20%
Project Archives	15%
Reference Resources (Guides to Good Practice, Reports, Analyses, etc.)	10%
Teaching and Learning Specific Resources (Data Tools, In-situ, Book etc.)	5%
Bibliographies (Literature, Architecture Group etc.)	2%

Education has a problem with an abundance of information much of which becomes lost or inaccessible

- Information is more accessible
- More written information
- More published
- Increase in grey literature
- Information overload

Houghton, J. (2012). The value of alternative forms of publishing for policy research. Presentation given at the *Where is the Evidence?* conference, June 10, 2012. Canberra, Australia.

Centre for Strategic Economic Studies

---

---

---


---

---

---

---

---



## Lost information

**Lost**  
**Why?**  
**Not archived**

Conference papers

Dissertations

Newsletters

Research projects

Policy statements

Annual reports

Working papers

White papers

Bulletins

Fact sheets

Blogs

Technical reports

Memoranda

Meeting minutes

Speeches

Statistics

Directories

Programs & projects

Memoranda

Guidelines

Surveys

Bibliographies

---

---

---

---


---

---

---

---





## Integrators/Curators

**Digital Education Research Network (DERN)**  
<http://dern2.acer.edu.au>

- Research reviews
- Research news
- Research briefs
- Research database
- Archives
- Email alerts

## Born digital

(OCLC)

---

---

---


---

---

---

---

---



## Tyndall framework

... to systematically build a search schema for a discipline

ACCORDS		
<i>Govt reports</i>	• Authority	<i>Open resources</i>
<i>Blogs</i>	• Accuracy	
<i>Research reports</i>	• Coverage	<i>National bodies</i>
<i>Projects</i>	• Objectivity	
<i>Conferences</i>	• Date	<i>International bodies</i>
<i>Newsletters</i>	• Significance	

Tyndall, J. (2008). *How low can you go?: Toward a hierarchy of grey literature*. Retrieved September 8, 2012, from <http://dspace.flinders.edu.au/dspace>.

---

---

---


---

---

---

---

---



## Gerry White

# Thank you

**Principal Research Fellow**  
[Gerry.White@acer.edu.au](mailto:Gerry.White@acer.edu.au)  
<http://dern2.acer.edu.au>  
**Digital Education Research Network**

---

---

---

---

---

---

---

---

## Research Life Cycle: Exploring Credibility of Metrics and Value in a New Era of eScholarship that Supports Grey Literature

**Julia Gelfand**, University of California, Irvine (UCI)

**Anthony Lin**, Irvine Valley College, United States

The fundamental components of the research process are defined by academic tradition, discipline and its participants. Traditional scholarship has now evolved into eScholarship with emerging technologies providing new methods of innovation and new ways of handling classical research processes. This revised research life cycle not only incorporates the established parts of the research chain, from discovery, gathering, and creating, but now has added phases of citing, sharing, preserving and archiving. There are quantifiable elements that help describe unique elements depending on sector and subject matter and format. Previously defined barriers such as geographical, institutional, digital and domain boundaries that previously existed can now be transcended to either accelerate or retard the research lifecycle in amazing and innovative ways. This new paradigm shift of current practices or activities today include the range of literacies that must be demonstrated and include information literacy, visual literacy, financial literacy, and increasingly data literacy. The role of the academic library has become increasingly visible as scholars and scientists seek support in managing their research lifecycle components. Librarians are now managers and curators of the scholarly research lifecycle by protecting, harvesting, and promoting reuse of content for new and unprecedented purposes. In a similar fashion, Grey Literature has previously followed the lifespan of more traditional output. Now new technologies exist to extract value metrics that compare favorably with other information products. This paper will explore how Grey Literature matures through different pathways or life cycles as the new grey becomes less grey with increasing value metrics to support and describe it. Also, the world of publishing has become increasingly accessible to a new population of scholars to release new information and ideas, contribute to emerging fields and frontiers without the barriers or requirements of following a specific trajectory of traditional publication processes. Examples will be shared about how the research life cycle has evolved with new tools to support Grey Literature from the life cycle management (LCM) and life cycle assessment (LCA) models to determine impacts and drive future directions concerning options for actions like open access, intellectual property and other forms of rights management.

### Bionotes

**Julia Gelfand** is the Applied Sciences and Engineering Librarian at the University of California, Irvine and has written and presented extensively for over two decades on different aspects of grey literature and its challenges in collection development for libraries and the emerging technologies that support grey literature.

Email: [jgelfand@uci.edu](mailto:jgelfand@uci.edu)

**Anthony Lin** formerly a Business Librarian, is currently the Head of Reference and Technical Services at the Irvine Valley College Library, Irvine, CA, USA and this is the 4th major presentation that these authors have collaborated on.

Email: [alin@ivc.edu](mailto:alin@ivc.edu)

## List of Participating Organizations

*Last Update: November 10, 2012*

Australian Council for Educational Research, ACER	Australia
Bangalore University	India
Battelle Memorial Institute	United States
Benaki Phytopathological Institute	Greece
Biblioteca Centrale "G. Marconi"; CNR	Italy
BMS College of Engineering	India
British Library, BL	United Kingdom
Centre National de Recherche Scientifique, CNRS	France
Dalhousie University, DAL	Canada
Data Archiving and Networked Services, DANS	Netherlands
EBSCO Publishing	United States
Elsevier	Netherlands
European Commission, EC	Belgium
European Food Safety Authority, EFSA	Italy
Federal Library Information Network, FEDLINK	United States
Food and Agriculture Organization of the United Nations, FAO	Italy
Government P.U. College	India
Grey Literature Network Service, GreyNet	Netherlands
Health Information Network Calgary, HINC	Canada
Information International Associates, Ila	United States
Institut de l'Information Scientifique et Technique, INIST	France
Institute of Information Science and Technologies, ISTI-CNR	Italy
Institute of Marine Sciences, ISMAR-CNR	Italy
Institute of Research on Population and Social Policies, IRPPS	Italy
Irvine Valley College Library	United States
Istituto di Linguistica Computazionale, ILC	Italy
Istituto Superiore di Sanità, ISS	Italy
Japan Atomic Energy Agency, JAEA	Japan
Japan Science and Technology Agency, JST	Japan
Korea Institute of Science & Technology Information, KISTI	Korea
Library of Congress, LC	United States
Mysore University Library	India
National Diet Library, NDL	Japan
National Documentation Center	Greece
National Research Council, CNR	Italy
National Technical Library, NTK	Czech Republic
New York Academy of Medicine, NYAM	United States
Open University	United Kingdom
PricewaterhouseCoopers, PwC	Netherlands
Research Centre on Scientific and Technical Information, CERIST	Algeria
Science & Technology Facilities Council, STFC	United Kingdom
Senate Library	Italy
Slovak Centre of Scientific and Technical Information, CVTISR	Slovak Republic
Springer Verlag GmbH	Germany
Swinburne University of Technology, SWIN	Australia
Technische Informationsbibliothek, TIB	Germany
TextRelease	Netherlands
Thomson Reuters	Italy
Université Charles de Gaulle Lille 3	France
Université Claude Bernard Lyon 1	France
University of Bergen, UiB	Norway
University of Calgary	Canada
University of California, Irvine Libraries, UCI	United States

Pre-Conference Announcement  
*GL Conference Series 1993-2013*

# Fifteenth International Conference on Grey Literature **The Grey Audit: A Field Assessment of Grey Literature**

Bratislava, Slovak Republic 2-3 December 2013



Slovak Centre of Scientific and Technical Information



CVTI SR

GL15 Program and Conference Bureau

**TextRelease**

Javastraat 194-HS, 1095 CP Amsterdam, The Netherlands

[www.textrelease.com](http://www.textrelease.com) • [conference@textrelease.com](mailto:conference@textrelease.com)

Tel/Fax +31-20-331.2420

## Index to Authors

### A-B

Aissa, E.	103
Assante, Massimiliano	120
Asserson, Anne	29
Bardi, Alessia	21
Biagioni, Stefania	48
Boukacem-Zeghmouri, Chérifa	27
Brase, Jan	14

### C-D

Cancedda, Flavia	15
Candela, Leonardo	120
Candlish, J.R.	63
Castelli, Donatella	120
Chalabi, Lydia	34
Chowdappa, N.	58
Cinini, Alessandra	104
Congiu, L.	103
Crowe, June	63
Cucurullo, Sebastiana	108
De Biagi, Luisa	15
De Castro, Paola	151
De Santo, Elizabeth M.	54
Di Cesare, Rosa	145
Dušková, Marta	133

### F-G

Farace, Dominic	68
Frantzen, Jerry	68
Fukazawa, Takeyasu	73
Galatis, Helen	157
Gelfand, Julia	162
Giannini, Silvia	145
Goggi, Sara	89
Gonda, Mayuki	73
Gribling, Armand	43

### H-I-J-K

Hardouveli, Despina	139
Hayakawa, Misa	73
Houssos, Nikos	139
Itabashi, Keizo	73
Jeffery, Keith G.	29
Kalentsits, Maria	43
Kieniewicz, Johanna	78
Kitsiou, Maria V.	79
Kravjar, Július	133

### L-M

La Bruzzo, Sandro	21
Lawrence, Amanda	157
Lee, Seon-Hee	37
Lefebvre, D.	103
Lin, Anthony	162
Lin, Yongtao	126
Luzi, Daniela	48

### M-N

MacDonald, Bertrum H.	54
Manghi, Paolo	21
Marzi, Claudia	85
Morais-Pires, Carlos	13
Morganti, C.	103
Nagaya, Shun	73
Nanakos, Chrysostomos	139
Newbold, Elizabeth	78
Nobile, Marianna	95, 145
Noorani, S.	103

### P-Q-R

Pagano, Pasquale	120
Pardelli, Gabriella	89
Pecoraro, Fabrizio	95
Pejšová, Petra	99
Picchi, Eugenio	104, 108
Pirro, P.	103
Prost, Hélène	57
Quigley, Kevin	54
Ramasesh, C.P.	58
Richardson, J.	103
Roubani, Alexandra	139
Ruggieri, Roberta	48

### S

Sarantopoulou, Ioanna	139
Sassi, Manuela	89
Sassolini, Eva	104, 108
Sbrulli, Stefano	104
Schiano, Elisabetta	48
Schöpfel, Joachim	57, 114
Soomai, Suzuette S.	54
Soumplis, Alexandros	139
Souvlidis, Vasileios	79
Stathopoulos, Panagiotis	139
Stathopoulou, Ioanna-Ourania	139
Stock, Christiane	68
Sudlow, Alan	78

### T-U-V

Thomas, Julian	157
Usha Devi, L.	58
Vaska, Marcus	126
Verloo, D.	103
Vyčítalová, Hana	99


### W-X-Y

Weldon, Paul	157
Wells, Peter G.	54
White, Gerald	157
Witkamp, Paula	68
Yonezawa, Minoru	73

## Publication Order Form

2003-

### FOURTEENTH INTERNATIONAL CONFERENCE ON GREY LITERATURE

<i>Forthcoming February 2013</i>	<i>No. of Copies</i>	<i>x</i>	<i>Amount in Euros</i>	<i>Subtotal</i>
GL14 CONFERENCE PROCEEDINGS - Printed Edition ISBN 978-90-77484-20-3 ISSN 1386-2316 <i>Postage and Handling excluded<sup>*)</sup></i>		x	99.00 =	€
GL14 CONFERENCE PROCEEDINGS - CD-Rom Edition ISBN 978-90-77484-20-3 ISSN 1386-2316 <i>Postage and Handling included</i>		x	99.00 =	€
GL14 Conference Proceedings - Online Edition ISBN 978-90-77484-20-3 ISSN 2211-7199 <i>Password Protected Access</i>		x	99.00 =	€
 <b>POSTAGE AND HANDLING PER PRINTED COPY<sup>*)</sup></b>				
Holland		x	5.00	€
Europe		x	10.00	€
Other		x	20.00	€
<b>TOTAL =</b>				€

#### Customer Information

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 "GL14 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.