Eighteenth International Conference on Grey Literature

Leveraging Diversity in Grey Literature

The New York Academy of Medicine, USA • November 28-29, 2016

Proceedings

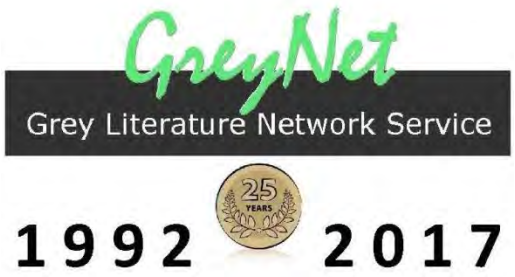
ISSN 1386-2316

Host and Sponsors

GL18 Program and Conference Bureau

TextRelease

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



CIP

GL18 Conference Proceedings
Eighteenth International Conference on Grey Literature : Leveraging Diversity in Grey Literature - New York, NY USA, on 28-29 November 2016 / compiled by D. Farace and J. Frantzen ; GreyNet International, Grey Literature Network Service. - Amsterdam : TextRelease, February 2017. -178 p. - Author Index (GL Conference Series, ISSN 1386-2316 ; No. 18).

The New York Academy of Medicine (USA), CVTISR (SK), DANS-KNAW (NL), EBSCO (USA), FEDLINK - Library of Congress (USA), Inist-CNRS (FR), ISTI-CNR (IT), KISTI (KR), NIS-IAEA (AT), NTK (CZ), and TIB (DE) are Corporate Authors and Associate Members of GreyNet International. These proceedings contain the full text of 16 conference papers presented during the two days of plenary and poster sessions. The papers appear in the same order as in the conference program book. Included is an author index with the names of some 40 contributing authors and researchers along with their biographical notes. A list of more than 50 participating organizations as well as sponsored advertisements are likewise included.

ISBN 978-90-77484-30-2
© TextRelease, 2017

Foreword

LEVERAGING DIVERSITY IN GREY LITERATURE

Scientific information, much of which is published as grey literature, can play a pivotal role in the search for solutions to global problems. Diversity invigorates problem solving and science benefits from a community that approaches problems in a variety of creative ways. Despite their diversity, the hundreds of authors and researchers across the globe involved in grey literature can be seen as part of the same community contributing to the scientific enterprise in valuable ways.

Diversity speaks directly to the effectiveness of information professionals working together as a team and is an essential ingredient for innovation. People from different backgrounds bring with them new information. If you want to build teams, communities, and organizations capable of innovating, you need diversity. It enhances creativity and encourages the search for new information and nuanced perspectives, leading to better decision making and problem solving. Diversity can improve the bottom line of companies as well as organizations, because exposure to it changes the way one thinks. A diverse community of researchers anticipate differences and understand that they will have to work harder to achieve consensus, but their diligence can lead to better outcomes. Authors in the GL-Conference Series come from different societal cultures and geographic regions; however in their research, they are united by the culture of science, which is without borders. This diverse community has over the past two decades applied research methods and offered explanations that have helped this field of information through blind spots, shedding light on what were once seen only as inherent problems. Their evidence based approaches have opened up new areas of research in grey literature. Where in the early '90s the focus was primarily on the demand side of grey literature, equal emphasis today is directed to its supply side. Speed and scale of communication are significant factors that contribute to diversity. The proliferation of technologies has allowed for an exponential growth of knowledge in information science just as in other sciences. However, the diverseness of grey literature resources has become a major challenge to its exploitation. The availability of systems for collecting and aggregating data and its semantic analysis has now become a priority.

GL18 focusses on evidence and seeks to further raise awareness among the wider public to the strength of grey literature based on a shared commitment by a diverse community of authors and researchers responsible for its production, open access, and digital preservation.

Dominic Farace — *Amsterdam,*

GREYNET INTERNATIONAL — FEBRUARY 2017

GL18 Conference Sponsors

The New York Academy of Medicine, USA

KISTI, Korea
Korea Institute of Science and Technology Information

CVTISR, Slovak Republic
Slovak Centre of Scientific and Technical Information

EBSCO, USA

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

TIB, Germany
German National Library of Science and Technology – Leibniz Information Centre for Science and Technology and University Library

GL18 Conference Sponsors

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

NTK, Czech Republic
National Library of Technology

FEDLINK, USA
Federal Library Information Network
Library of Congress

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

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

GL18 Program Committee

Danielle Aloia Chair

The New York Academy
of Medicine, NYAM
United States

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

Meg Tulloch
FEDLINK
Library of Congress
United States

Marnix van Berchum
Data Archiving and
Networked Services,
DANS-KNAW, Netherlands

Dobrica Savić
Nuclear Information
Section International
Atomic Energy Agency,
Austria

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

Petra Pejšová
National Library of
Technology
Czech Republic

Stefania Biagioni
Institute of Information
Science and Technologies
CNR, Italy

Deborah Rabina
Pratt Institute,
School of Information
United States

YongJu Shin
Korea Institute of Science
and Technology
Information, KISTI Korea

Joachim Schöpfel
University of Lille
France

Dominic Farace
GreyNet International
Grey Literature Network
Service, Netherlands

Tomas A. Lipinski
University of Wisconsin,
Milwaukee, UWM
United States

Table of Contents

Foreword 3

Conference Sponsors 4

Program Committee 6

Conference Chair and Moderators 8

Conference Program 9

Program

Session One – Diversity in Publishing and Disseminating Grey Literature 11

Session Two – Open Access and Shared Grey Resources 59

Panel Session – Data: Its Use and Preservation 91

Poster Session – Leveraging Diversity in Grey Literature 97

Session Three – Altmetrics, Aggregating Data and Semantic Analysis 117

Session Four – Fueling Public Awareness to Grey Literature 151

Advertisements

PsycEXTRA EBSCO 10

INIS, The International Nuclear Information System 58

FEDLINK, The Federal Library and Information Network - Library of Congress 70

NTK, National Library of Technology, Czech Republic 96

CNR, Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", Italy 116

The New York Academy of Medicine 150

KISTI, Korea Institute of Science and Technology Information 158

CVTISR, Slovak Centre of Scientific and Technical Information 176

Appendices

List of Participating Organizations 169

GL19 Conference Announcement 170

GL19 Call for Papers 171

Author information 172

Index to Authors 177

GL18 Publication Order Form 178

Moderator Day One

Meg Tulloch
Executive Director FEDLINK

Library of Congress

Meg is Executive Director of the Federal Library and Information Network (FEDLINK), Library of Congress. She is the former Library Director of the National Defense University Libraries in Washington, D.C. and Norfolk, Virginia. Previously, she was the Europe Region Librarian for the U.S. Army and oversaw 26 libraries in four different countries. She has also worked as a librarian at Vanderbilt University's Walker Management Library and Kutztown University of Pennsylvania's Rohrbach Library. Much of her career has focused on how technology can assist the researcher through digital library tools, using digital materials. Additionally, Meg taught "Introduction to Poetry Writing" at the University of Virginia while a graduate student there. She holds a Masters in Library and Information Science, a Masters in Fine Arts in poetry writing, and a Bachelors in American Literature. She is currently pursuing a Doctorate of Liberal Studies from Georgetown University. Her dissertation will explore fragmented twenty-first century literature.

Email: mtulloch@loc.gov

Program Chair

Danielle Aloia
Special Projects Librarian

New York Academy of Medicine

Daniella is Special Projects Librarian at The New York Academy of Medicine. She received her MSLS from Catholic University of America, in Washington, DC, in 2005 while working on the AgeLine Database at AARP. She has over 20 years of experience in a variety of library settings, including academic, non-profit and museum. She has been involved with collecting, evaluating, and cataloging grey literature since 2006, first at AARP and then at the United States Department of Transportation. For the past 4 years she has been managing the Grey Literature Report in Public Health, produced by NYAM.

Email: daloia@nyam.org

Moderator Day Two

Davis Erin Anderson
Community Engagement Manager

Metropolitan NY Library Council

Davis is Community Engagement Manager at the Metropolitan New York Library Council, where she helps make METRO a fun and friendly place to learn and try new things. Her work primarily focuses on bringing the library and tech communities together to find innovative ways to bring solid knowledge-building skills to the wider NYC citizenry. Davis received her Master of Science in Information and Library Science from Pratt Institute in 2013, a Master of Music from Western Michigan University in 2005, and a Bachelor of Music from The University of Wisconsin at Madison in 2003. Davis is super proud to have received SLA's Rising Star Award, and she is honored to have been included in Library Journal's Movers & Shakers Class of 2012.

Email: deanderson@metro.org

SESSION ONE – DIVERSITY IN PUBLISHING AND DISSEMINATING GREY LITERATURE

Collecting, organizing, and preserving diverse publication sources for the good of one community archive: Legal challenges and recommendations **11**
Andrea Copeland, Indiana University at Indianapolis; Tomas Lipinski, University of Wisconsin at Milwaukee; Kyle Jones, Indiana University at Indianapolis, USA

The recent improvements on circulation of research results at the Japan Atomic Energy Agency **24**
Satomi Inagaki, Misa Hayakawa, Naomi Ebisawa, Mayuki Gonda, Takashi Nozawa, and Keizo Itabashi, JAEA, Japan

Sexuality Leveraged Through Diversity: Recognizing LGBT+ Communities **32**
Anthony Lin, Irvine Valley College Library and Julia Gelfand, University of California, Irvine Libraries, USA

Managing Diversity in the International Nuclear Information System **49**
Dobrica Savić, Nuclear Information Section, International Atomic Energy Agency, Austria

SESSION TWO: OPEN ACCESS AND SHARED GREY RESOURCES

Looking for Information that is Not Easy to Find: An Inventory of LibGuides in Canadian Post-Secondary Institutions Devoted to Grey Literature **59**
Marcus Vaska, Alberta Health Services and Rosvita Vaska, University of Calgary, Canada

Bridging the Knowledge Gaps: What Grey Literature does for the Library of Congress' Indigenous Law Portal **71**
Carla Davis-Castro, Library of Congress, USA

Transition to Open Access and its implications on Grey Literature Resources **81**
Snježana Ćirković, Austrian Academic Library Consortium, Austria

PANEL SESSION – DATA: ITS USE AND PRESERVATION

Developing, linking, and providing access to supplemental genetics dataset vcf files **91**
Plato L. Smith II and Lauren McIntyre, University of Florida, USA

POSTER SESSION

Policy Development for Grey Literature Resources: An Assessment of the Pisa Declaration **97**
Dobrica Savić, NIS-IAEA, Austria; Dominic Farace and Jerry Frantzen, GreyNet International, Netherlands; Stefania Biagioni and Carlo Carlesi, ISTI-CNR, Italy; Herbert Gruttemeier, Christiane Stock, Inist-CNRS, France

A Geographical Visualization of GL Communities: A Snapshot **109**
Gabriella Pardelli, Sara Goggi, Roberto Bartolini, Irene Russo, and Monica Monachini, Istituto di Linguistica Computazionale "A . Zampolli" CNR Pisa, Italy

Teaching and Learning about Grey Literature : Results from a Poster Presented at the 18th Grey Literature Conference **114**
Lynne Rudasill, Center for Global Studies; University of Illinois, United States

SESSION THREE – ALTMETRICS, AGGREGATING DATA, AND SEMANTIC ANALYSIS

A terminological "journey" in the Grey Literature domain **117**
Roberto Bartolini, Gabriella Pardelli, and Sara Goggi, ILC-CNR, and Silvia Giannini, Stefania Biagioni, ISTI-CNR, Italy

Altmetrics and Grey Literature: Perspectives and Challenges **131**
Joachim Schöpfel, GERiiCO Laboratory, University of Lille and Hélène Prost, CNRS France

SESSION FOUR – FUELING PUBLIC AWARENESS TO GREY LITERATURE

The GreyLit Report: Understanding the Challenges of Finding Grey Literature **151**
Danielle Aloia and Robin Naughton, The New York Academy of Medicine Library, USA

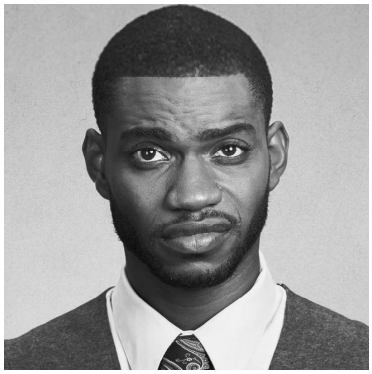
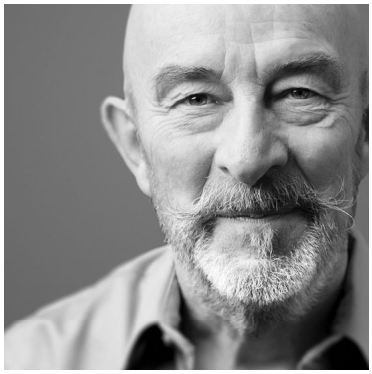
Debate about Scientific Popularization in Russian Public Sphere (Based on Grey Literature Material) **159**
Yuliya B. Balashova, Saint Petersburg State University, Russia

'Grey crossroads' in cultural heritage preservation and resource management **162**
Luisa De Biagi, CNR Central Library and Roberto Puccinelli, Telecommunications & Informative Systems Office, Italy

PsycEXTRA®

EXPAND BEHAVIORAL SCIENCE RESEARCH WITH THE PREMIER RESOURCE FOR GRAY LITERATURE

Devoted to curating and indexing hard-to-find content from authoritative sources, PsycEXTRA® allows researchers to go beyond traditional peer-reviewed research materials. This high-quality and relevant database combines bibliographic records with unique full-text materials, focusing on the latest conference presentations and papers, newsletters, reports, patient-oriented factsheets and brochures, magazines, monographs, and standards and guidelines relevant to the needs of students, faculty, and clinicians alike.


Explore original, cutting-edge, high-quality research

- Updated biweekly
- More than 300,000 records and growing
- No coverage overlap with PsycINFO®, creating an ideal companion database
- Full text for more than 70% of records
- Ongoing updates of select professional literature from multiple state and government entities, associations, foundations, and more
- Average of 25,000 records or more added each year

Available via
EBSCO

Collecting, organizing, and preserving diverse publication sources for the good of one community archive: Legal challenges and recommendations

Andrea Copeland, Indiana University at Indianapolis
Tomas Lipinski, University of Wisconsin at Milwaukee
Kyle Jones, Indiana University at Indianapolis, USA

Over the past several years, the bicycle movement in Indianapolis, Indiana, has gained a great deal of momentum. Bicycle lanes and trails have been designed and implemented to support commuting to work, rental stations have been installed in strategic locations across the city, and a bike hub has been built providing showers, lockers, and repair services. The 2007-2015 mayoral administration was a strong advocate for the movement, and the city has received over $60 million in private donations and federal grants to build the Indianapolis Cultural Trail, which is at the heart of the improved pedestrian and cycling infrastructure. The city has also developed 74 miles of on-street bike lanes within the metropolitan area. Representatives from other national and international municipalities have visited the city to study Indianapolis's cycling infrastructure. These changes have brought about economic growth, improvement to the general air quality, and personal health gains.[1]

Despite the extensive investment in fostering a culture of cycling, there is not yet a significant formal mechanism for documenting or analyzing the effects of these changes. A future goal of this research is to develop a community archive development processes/prototype application that will provide such a tool. Indianapolis is an ideal city for the system design as cycling infrastructure in Indianapolis is emerging and developing rapidly. continuum of cycling infrastructure development.

The prototype (archive and mobile application) will be called CHIME: Citizen-data Harvest in Motion–Everywhere. CHIME will be designed to combine the data collected in the form of the various community voices expressed in a diversity of formats (images, video, textual narratives, and geographical information), along with overlays of geographic data to provide context.

The cycling experience in Indianapolis is an ideal community issue for which to develop these processes, as the geographic and mobile nature of the phenomenon will expose the challenges of capturing both place-bound and digital history as it is happening. Most information regarding cycling is written on the landscape or in a digital form. Much like changes to the physical landscape of a city, current digital information can be difficult to grasp all at once as it is widely distributed. In this way, cycling produces both tangible and intangible cultural artifacts and provides a venue for exploring the preservation and sharing of both types of artifacts.

The perception that communities are a harmonious whole is seldom accurate. Tensions of varying degrees between different groups invariably exist. In some cases, the group divisions are based on obvious distinctions such as religious preferences, political beliefs, income, and ethnicity. In the case of the bicycle movement, these differences serve as a backdrop to the common dilemma of how to best use shared public space. In the case of Indianapolis, not all residents are supportive of the movement. Most of the dissent manifests in community news forums and in broadcast and print media. Motorists have been vocal about not wanting to share the roads. Residents in predominantly African-American neighborhoods are upset that the city chose to equip roadways in their communities with bike lanes without asking for their input.[2]

[1] Sustainable Communities. (2013). Indianapolis Cultural Trail: Improving livability in Central Indiana. www.sustainablecommunities.gov/sites/sustainablecommunities.gov/files/docs/2013_5_23_indianapolis_case_study.pdf

[2] Brown, A. (2012). Indy's bike lanes blasted by black community; Question why, the cost and the need. PraiseIndy.com. http://tinyurl.com/m6ugk66

Humanities-based practices of public scholarship and civic engagement are particularly important to the design of the prototype. Scholars in fields such as cultural studies, public history, digital humanities, and museum studies have long recognized that archives, exhibitions, and official histories have privileged the perspectives and experiences of people and institutions in positions of power. While city records and newspaper reports (among other sources) are already documenting the ways in which policy makers and commercial entities are engaging with the changes in the city's cycling infrastructure, it is harder to collect, preserve, and make known the experiences of individual cyclists, drivers, and pedestrians. For example, the public discourse taking place in the comments sections of blogs and online newspaper articles, as well as the personal snapshots and reflections published via social media platforms, are of a troublingly ephemeral nature. If we are to document a community's experience, we must look beyond official accounts and acknowledge the ways in which individuals' daily lives shape and are shaped by cycling infrastructure. Capturing a diverse array of perspectives is just one part of the task.

Even more challenging than incorporating multiple perspectives is the task of capturing individuals' experiences of space while they inhabit it. Life events are merged with objects, buildings, and places; one is not realized independent of the other.[3] To better document events that occur within a community space, the corresponding human experience needs to be captured simultaneously. This is a challenging proposition, especially when one considers the growing influence of tools and interaction in the digital realm on the human narrative.

Documentation and dialogue captured or facilitated with digital tools and platforms is our collective social history and resides mainly in the hands of corporate entities like Facebook, which could cease to exist tomorrow or choose to shut down platforms without notice. Further, corporations have no legal obligation to preserve our collective social heritage and have often demonstrated a lack of concern for users and user-created content.[4] For the most part, heritage organizations are not involved in the process of collecting and preserving born digital information generated in this everyday digital context from individuals or communities of individuals. This paper will explore the legal and ethical issues involved with the capture of the kinds of data (see Table 1 for list of data types) described in support of building the prototype archives.

Table 1. Community Archive (case study CHIME, Indianapolis, IN): Legal Issues

Document	***Copyright and Licensing Other Issues***	***Privacy***	***Zoning and Related Regulation***	***Public Domain***
Photographs – personal, community events Videos (e.g. GoPro)	Coprightability: pictorial works. *Bridgeman Art Library, Ltd. v. Corel Corp.*, 36 F.Supp.2d 191 (S.D.N.Y. 1999). Thin Copyright and fair use. Ownership: Work made for hire.	Intrusion: intrusion and appropriation.		
Photographs shared via social media platforms	CMI (17 U.S.C. § 1202): *Cable v. Agence France Presse*, 728 F.Supp.2d 977 (N.D.Ill. 2010) ("Cable created [] works ... reproduced … and attributed [] as follows 'Photos ©2009 wayne cable, selfmadephoto.com.' …." Id. at 978. "[U]nder the plain language of the statute [] plaintiff's name and hotlink fall within the scope of 'copyright management information.'" Id. at 981. Ownership and use: End User License Agreement. See below.	Privacy: intrusion and appropriation.		
Videos shared via Social Media	Ownership and Use: End User License Agreement. Some social media websites claim either ownership or non-exclusive rights to use content posted photographs so posted. See, Lipinski and Copeland (2013).	Privacy: intrusion and appropriation.		

[3] Halbwachs, M. Space and the collective memory. In *The Collective Memory*; Harper & Row: New York, NY, 1950

[4] http://kernelmag.dailydot.com/issue-sections/features-issue-sections/16616/archive-team-saving-the-web/

Document	***Copyright and Licensing Other Issues***	***Privacy***	***Zoning and Related Regulation***	***Public Domain***
Online News Articles (e.g. Indy Star)	Thin copyright and fair use. Literary works. See, *Los Angeles Times v. Free Republic*, 29 Media L. Rep. 1028 (C.D. Cal. 2000) (online discussion board use of complete articles not fair use). Ownership: Work made for hire. Source of clippings: End User License Agreement?			
Comments contributed at the end of the articles (often document the conflict generated by motorists not wanting to share the road)	Thin copyright and fair use of literary works. Ownership and Use: End User License Agreement. Defamatory of otherwise tortious: application of 47 U.S.C. § 230?			
Newsletters of these advocacy groups	Fair use: Congress has recognized that "the scope of the fair use doctrine should be considerably narrower in the case of newsletters than in that of either mass-circulation periodicals or scientific journals….newsletters are particularly vulnerable to mass photocopying, and ... most newsletters have fairly modest circulations." H.R. No. 94–1476 at 73 (1976), *reprinted in* 1976 U.S.C.C.A.N. 5659, 5687.			
Social media and websites cycling advocacy groups Include government documents: city reports, crime (bike theft) and accident data, routes, maps, city planning documents, minutes from the Mayors Advisory Council on Cycling	States and local governments are free to claim copyright protection subject to the standards of the copyright law. For example, the crime data are not protected by copyright as purely factual content unless selected, coordinated or arranged in some creative manner as a compilation copyright. Routes/Maps: The question would be whether "the overall manner in which [the plaintiff] selected, coordinated, and arranged the expressive elements in its map, including color, to depict the map's factual content." *Streetwise Maps, Inc. v. Vandam, Inc.*, 159 F.3d 739, 748 (2d Cir. 1998).			Works of the federal government are in the public domain. 17 U.S.C. § 105. Public policy dictates that the documentary building blocks of law and government; its cases, statutes, regulations, etc. are in the public domain: "It is well settled that judicial opinions and statutes are in the public domain and are not subject to copyright." *Veeck v. Southern Bldg. Code Congress Intern. Inc.*, 241 F.3d 398, 412 (5th Cir. 2001).
TV news Radio shows discussing cycling	Thin copyright and fair use of audiovisual works. "For the foregoing reasons, TVEyes' archiving function qualifies as fair use, and its downloading and Date–Time search functions do not qualify as fair use. Its e-mailing feature can qualify as fair use, but only if TVEyes develops and implements adequate protective measures." *Fox News Network, LLC TVEyes, Inc.*, 2015 WL 5025274, *10 (S.D. N.Y.). *Maxtone-Graham v. Burtchaell*, 803 F.2d 1253 (2d Cir. 1986), cert. denied 481 U.S. 1059 (1987). Ownership: Work made for hire.			
Comments on TV news clips posted to station website	See above, same analysis as commentary on news articles.			

Document	***Copyright and Licensing Other Issues***	***Privacy***	***Zoning and Related Regulation***	***Public Domain***
Blogs* Example: Neil Kelty blog of his commuting by bike death in a traffic accident with a school bus: https://medium.com/urban-cycling/a-novice-cycles-to-work-ad5df66c937#.jvyajrtwz	See above regarding specific content and circumstances: photographs, audiovisual works and commentary, etc. Blog posts may not be not be protected by its state shield laws as are traditional news reporters and publishers: "By contrast, defendant's comments on an online message board would resemble a ***pamphlet*** full of ***unfiltered, unscreened letters*** to the editor submitted for publication—or, in modern-day terms, ***unedited, unscreened comments posted by readers*** on NJ.com." *Too Much Media, LLC v. Hale*, 993 A.2d 845, 847 (N.J. Superior Court 2010), affirmed 20 A.3d 364, 379 (N.J. 2011).			
Ghost Bikes These bikes mark where someone was killed, see http://ghostbikes.org				Roadways, sidewalks, bike paths and other public right of ways are government property. Regulation is subject to constitutional requirements: regulation should be content neutral or strict scrutiny analysis applies. Content neutral, reasonable time, place and manner restrictions must be viewpoint neutral.
Street Art decorating city pedestrian and cycling path ways: http://www.indianatrails.com/content/trails-and-public-art	Fair use of pictorial work: *Seltzer v. Green Day, Inc.*, 725 F.3d 1170, 1173 (9th Cir. 2013): "Staub photographed a brick wall at the corner of Sunset Boulevard and Gardner Avenue in Los Angeles which was covered in graffiti and posters—including a weathered and torn copy of *Scream Icon*." Id.			
Databases – Routes, trails, uploaded, MapMyRide and Garmin – using GPS technology.				*County of Santa Clara v. Superior Court*, 89 Cal.Rptr.3d 374, 3934 (Cal. App. Dist. 6, 2009): "Matching the GIS Basemap with orthophotographs, which are in the public domain…"
Community Cycling events – Tweed Ride, NITE Ride (etc).	See above regarding specific content and circumstances: photographs, interviews and commentary and interviews, audiovisual works, etc.			See above regarding specific content and circumstances: route map and geospatial information.

Ethical Issues Explored

CHIME has positive aims. Through data and technology, it seeks to create a socio-technical infrastructure that can speak to the experiences and interests of a broad array of stakeholders. In so doing, it will become an empowering technology that, among other things, will give individuals and groups the opportunity to have a voice in the way their city is molded and captured in history. Regardless of the good intentions, there are a number of important ethical considerations to address.

First, there is a question of whether or not the burden of documentation is justifiable given possible invasions of privacy. The public will be subject to archived documentation in perpetuity. And while documentation born from public places and spaces seems defensible

from a legal standpoint, it may be that CHIME administrators have an ethical duty to provide for fair information practices should members of the public feel that their inclusion in the archive is a personal invasion of privacy.

Second, it is unclear at this point who will receive the benefits of CHIME's intended goods. On the face of it, CHIME's administrators position themselves as neutral entities, as system designers and researchers simply providing information resources for its users. But all data is "cooked" in some way, and the privileged position of technological designers always needs to be interrogated for embedded values and interests that may disenfranchise other parties.

Third, the data aggregated for and analyzed as a result of CHIME holds the potential to create significant value and create tensions around data ownership. Actors with the right technical skill set or in a position of power (e.g., city and university administrators) may claim that the datasets and resulting algorithms are a new product over which they can claim ownership and subsequently monetize. In contrast, the public has an equally powerful ownership claim on the grounds that such information products would not exist were it not for their active or passive participation in the project to begin with.

Finally, there is an open problem concerning who is or should be held responsible for CHIME's maintenance. Building an open data and community-based infrastructure like CHIME aims to do requires significant financial expenditures, and the labor involved will be highly specialized. While the project is oriented towards community needs, it would be unfair to place burden of its maintenance and future development on public volunteers.

In what follows, we briefly consider each of these four thematic ethical issues. None of the discussion definitively answers or resolves the problems; however, it lays an important foundation on which to examine how CHIME may bring to the fore moral problems related to rights, responsibilities, and benefits and burdens.

The Burdens of Documentation

CHIME seeks to aggregate a wide variety of data and information as part of a larger initiative to document and analyze Indianapolis' cycling infrastructure. The legal analysis of data sources suggests that, in most cases, information observed in public places and extracted from online sources does not trigger any specific individual right to privacy. For instance, photographs taken of cyclists on public roads or trails do not intrude in private spheres of life. Regardless of whether or not there exists a legal protection against intrusion, the CHIME project's documentation may burden members of the public. Three specific problems exist with respect to documentation: 1) the perpetuity of that which is documented, 2) the decontextualization of documentation, and 3) the ways in which documentation subjects can (or cannot) express agency.

Only through a longitudinal perspective will CHIME's primarily goals be met. That is, CHIME aims to document the changes in cycling landscape over time, which requires the ongoing documentation of cycling experiences expressed in data and information. It is imaginable that subjects caught in the data net CHIME administrators throw onto the public would have valid arguments for not wanting their cycling life archived in perpetuity, even though they might night have legal standing. Cycling can be a social experience shared between family members, friends, and acquaintances. And it is entirely plausible that individuals associating with each other on a public cycling path believe that their ride is ephemeral, that it leaves no lasting history. Now, the same individuals may willingly submit themselves to surveillant gaze of the city for the purposes of participating in a "safe cycling" campaign where police have access to trail-based cameras. But those individuals might think differently about having their association captured in a publicly available database.

Consider, also, information mined from a community Facebook group of cyclists. In this space, the members freely share their routes, experiences, and opinions on the state of cycling in their neighborhood. The information is contextualized, there are norms of reciprocity, and members develop trustworthy friendships with one another, bonding over the cycling experience and in so doing create a willingness to exchange thoughts and ideas.

Decontextualizing the information the members share and depositing it in a public archive for analysis and wider consumption is a prima facie threat to contextual integrity.[5] In other words, the migration of the information into another context immediately raises privacy issues–in the social but not the legal sense–because normative expectations about how that information should be accessed, used, and disseminated are no longer respected.

It is not to CHIME's advantage to create mechanisms by which particular sets of public information (e.g., videos capturing public individuals) can be suppressed or expunged from the archive. The more data the CHIME administrators and researchers have at their disposal, the argument goes, the wider and greater the insights they can distill from the archive. All that aside, harms the could accrue from perpetual documentation and decontextualization of public information may be manageable through fair information practices built into CHIME's information policies and technological infrastructure. Such practices may include the ability for an individual to express a privacy harm, prove her identity in relationship to CHIME's documentation, and request suppression of her identity or that the documentation be removed from the archive. By following this approach, CHIME would respect the interests of concerned individual's included in its archive without necessarily devaluing analytic findings.

Balance of Benefits

CHIME is, first and foremost, a research endeavor. Its creators aim to develop tools and insights that push the boundaries of knowledge. This requires the researchers to develop and disseminate their research in journal articles and presentations. Thus, the initial benefits of the project will redound to CHIME researchers who will reap professional and scholarly goods. But what about the data subjects who inform the project with their social media? Or the partners in city administration who provide insights into the cycling infrastructure? Or the general public who are captured in videos of cycling routes? Given that the scope of participants in the CHIME project is large, CHIME must carefully consider if its benefits will be equitably distributed.

The principle of beneficence in research maintains that the welfare of the research participant should be a guiding goal; however, it is challenging for researchers to claim who will benefit from their research and when those benefits will accrue. These types of moral calculations are always fraught with unknown variables. Regardless, researchers are burdened with conceptually mapping how their interventions or programs will create foreseeable benefits in the short and longterm and for whom.

CHIME aims to create a community platform of data and information for the public to access. While CHIME require users of the data to create accounts and agree to a terms of service, researchers will not be able to know exactly how the public will use the data and to what ends–good or bad. On one hand, it could be that a CHIME user analyzes the data to peddle erroneous information to heavily-biked areas about future city planning projects. Surely, this would be a negative use of the data in that the data is not being used to promote positive benefits in the community. On the other hand, community artists might use the data to develop place-based art to beautify heavily-biked areas and engage populated spaces. Many in the public would agree that this is a good use of the data. In both cases, however, these uses would be unknowable to CHIME's creators.

To develop equitable benefits, CHIME must strategically develop objectives that map to specific research participant groups. It is not enough to release the data into the wild via the community and hope for benefits to result and outweigh potential harms. To these ends, CHIME can develop community engagement strategies for, among others, bicyclist advocacy groups, city administrators, neighborhood businesses, and the like to share how to use CHIME and be specific about how CHIME might benefit them.

[5] Nissenbaum, H. (2010) *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Palo Alto, CA: Stanford University Press.

Data Ownership

Users of social media, like Facebook and Twitter, agree to end user license agreements (EULAs) or terms of service (TOS) upon registering for their accounts. EULAs and TOS legally bind users to a contract dictating, among other things, the rights the service provider and the use retain related to the data and information (broadly defined as content) created while interacting with the social media site. Yet, it is not clear if CHIME's action, such as aggregating social media into its databases, obligates researchers to respect TOS and EULAs users previously agreed to.

Facebook grants itself the right to use a user's data in its TOS. The company writes that its users give them "a non-exclusive, transferable, sub-licensable, royalty-free, worldwide license to use any IP content that [users] post on or in connection with Facebook" (https://www.facebook.com/terms). Moreover, users who post information publicly without using privacy restrictions "means that [users] are allowing everyone, including people off of Facebook, to access and use that information, and to associate it with [the content creator]." So, Facebook is at liberty to disclose user data and, at the same time, places no restrictions on third parties, such as CHIME, to access and manipulate user data, as long as it is public.

Twitter's TOS provides similar rights to itself, in that it retains the ability to distribute user data to third parties. For instance, Twitter can make user content available through APIs, business platforms, and by other technical routes, as long as individuals who gain access to user content agree to the terms and conditions governing content use. Where CHIME is concerned, this TOS provides less freedom to scrape publicly available sites–like Facebook groups or accounts–by requiring researchers to use Twitter approved methods to aggregate its content. Moreover, CHIME cannot distribute Twitter content to other parties via its own API: it can only distribute tweet or user IDs. It can, however, provide downloadable spreadsheets or PDFs of user content with some restrictions (https://dev.twitter.com/overview/terms/policy.html).

The Facebook and Twitter cases represent the complexity CHIME faces in not only accessing user data but also distributing it via its community platform. But, more importantly it shines light on the complicated matter of data ownership. Twitter explicitly states that users own their content, yet empowers its business partners to make use of user content for remunerative purposes without active user consent. Similarly research projects, like CHIME, can make derivative datasets of user data, again, without informing users. This is a loose and fast definition of 'data ownership."

While CHIME is not in a position to force social media companies to rethink their data ownership definitions and related policies, it should justify its position on the manner transparently to be in the clear, ethically speaking. CHIME should create information policy that discusses how it has gained access to social media content and why it has a right to do so. This policy should also state what rights CHIME has to the data as the curator of the dataset; similarly, the policy should state what rights data subjects retain. As the TOS analysis above shows, 'data ownership' is a legally complex concept. CHIME would benefit by not expressing its right to own data acquired by public means, but rather by expressing its role as a steward over the data.

Infrastructure Maintenance

CHIME is an advanced technological system that will undoubtedly require maintenance in order to maximize the informational and social goods it seeks to produce. Its database structures archive the useful data and information; its application programming interfaces (APIs) enable data access; its algorithms help to analyze the data; and its interfaces make the data and information usable for the community. All of these technical components require an advanced technical skill set to maintain into the future, which requires significant funding to pay for the labor. While CHIME may be sustainable in the near term with sufficient grant funding, its success in the future is still unknown. The ethical question here concerns who should be held responsible for the upkeep of the infrastructure. Put a different way and more specifically, we might ask who is morally obligated to maintain all of the technology once the initial funding runs out.

Only after the initial research is done will stakeholders, like community members, local businesses, and the city's administration, will be able to use the technological infrastructure for their own ends. But if the project is not sustainable in the long-term due to maintenance issues, it is less sure that these stakeholders will be able to use CHIME to, among other things, meaningfully support city planning projects or promote neighborhood cycling needs. In this scenario, CHIME's creators benefit from the data and information provided by the public without returning anything back to those who have supported their project.

To account for this issue and work towards the overarching goal of creating a sustainable archive, CHIME researchers have a responsibility to plan for extending the infrastructure's life and establishing end-of-life circumstances. Optimally, the researchers themselves will be able to find new funds or integrate CHIME into a university's technical infrastructure. But should this option not materialize, researchers need to account for the costs of CHIME's maintenance and the skill sets required to maintain for a certain period of time. Doing so will help the researcher's communicate to potential partners, like the city of Indianapolis, what resources are required to maintain CHIME. Moreover, the researchers should consider developing CHIME with sustainability in mind. This may require the researchers to think about and subsequently design for a CHIME-lite version without the technical 'bells and whistles.'

Legal Issues

The creation, preservation and use of the documentary record of the biking community of Indianapolis raise several issues including copyright and licensing, privacy and regulation of speech. Many of the copyright and licensing issues were covered in previous GL conferences and publications and are not discussed here.[6]

However, the copyrightability of local municipality official records (proceedings, minutes, etc.) and data, in specific geographic information raises new issues and are discussed in addition to the privacy (intrusion and appropriation rights) and First Amendment concerns raised by regulation of ghost bike sites.

Regulation of Ghost Bikes and other Memorials: Reasonable Time Place and Manner Restrictions

Some states and municipalities prohibited the erection of road side or curbside monuments or memorials. In the alternative some states or municipalities may allow be require removal after some time period such as thirty days. Such regulations are often a combination of state and local authority control as a municipality may have state roadways traversing its boundaries.

[6] Tomas A. Lipinski and Katie Chamberlain Kritikos, Copyright Reform and the Library and Patron Use of Non-text or Mixed-Test Grey Literature in Digital Scholarship, 12 THE GREY JOURNAL, INTERNATIONAL JOURNAL ON GREY LITERATURE (Grey Publishing, Licensing, and Open Access), No. 2, at 67 (2016), selected by the editors (pp. 67-81).
Tomas A. Lipinski and Andrea J. Copeland, Is the Licensing of Grey Literature
Using the Full Palette of "Contractual" Colors?, 11 THE GREY JOURNAL,
INTERNATIONAL JOURNAL ON GREY LITERATURE (Grey Publishing, Licensing, and Open Access), No. 2, at 69 (2015), selected by the editors (pp. 69-87).
Tomas A. Lipinski and Andrea Copeland, Look before you License: The Use of
Public Sharing Websites in building Patron Initiated Public Library Repositories, 42(4)
PRESERVATION, DIGITAL TECHNOLOGY & CULTURE, at 174, November, 2013
(Vol. 42, No. 3, pp. 174–198). ISSN (Online) 2195-2965, ISSN (Print) 2195-2957, DOI: 10.1515/pdtc-2013-0028.
Joachim Schöpfel and Tomas A. Lipinski, Legal Aspects of Grey Literature, THE
GREY JOURNAL: INTERNATIONAL JOURNAL ON GREY LITERATURE
(Managing Change in Grey Literature), Autumn, 2012 (Issue 8, No. 3) (pp. 137-153) (ISSN 1574-1796, available at http://www.greynet.org/thegreyjournal/previousissues.html).
Tomas A. Lipinsk, Green Light for Grey Literature? Orphan Works, Web-Archiving and other Digitization Initiative—Recent Developments in U.S. Copyright Law and
Policy, THE GREY JOURNAL: AN INTERNATIONAL JOURNAL ON GREY
LITERATURE, Spring 2009, at 11 (11-21), also available at http://www.greynet.org/thegreyjournal.html (ISSN 1574-1796).

Indiana House Bill 1108 was introduced in 2009 to require that "uniform roadside memorials" be erected at the request of the "immediate family" with erection lasting no more than one year. Proposed Indiana Code § 8-23-5-10(a)(b) and (c). The statute would have required the Indiana Department of Transportation to "remove roadside memorials that are not erected or sanctioned by the department." Proposed Indiana Code § 8-23-5-109(e). A "unit" other than the Indiana Department of Transportation could erect a "uniform roadside memorials within the right-of-way alongside highways, street, or roads within the unit's jurisdiction" but the erection must comply with the above requirements, i.e., Proposed Indiana Code § 8-23-5-10. Proposed Indiana Code § 36-1-4-6.5. The bill did not pass.

Another example is found in Milwaukee proposed Ordinance § 116-7 regulating "Roadside Memorials" defined as "any of the following items including but not limited to balloons, flowers, pictures, stuffed animals and religious items commemorating the site of a fatal accident or occurrence." Milwaukee proposed Ordinance § 116-7(1). The ordinance prohibits such memorials "on or within the boundaries of any public street, sidewalk or walkway for more than 14 days from the date of a fatal accident or occurrence." Milwaukee proposed Ordinance § 116-7(2). Here the triggering date is not when the memorial was first erected but from the date of the accident the memorial is intended to commemorate. The Milwaukee ordinance requires the commissioner of public works to "dismantle and discard a roadside memorial left on or within the boundaries of any public street, sidewalk or walkway after 30 days from the date of a fatal accident or occurrence." Milwaukee proposed Ordinance § 116-7(3). City of Milwaukee, Legislative Reference Bureau File Number: 050770 (March 23, 2006).

Milwaukee City Attorney Grant F. Langley indicated that a municipality could not allow one type of sign or message or memorial but not others. This could constitute impermissible viewpoint discrimination under the First Amendment of the U.S. Constitution. Reasonable time, place and manner restrictions that do not target a specific message or speaker are allowed. Duration, e.g., 30 or 90 days, location, e.g., on the curbside but not the medium strip, not higher than 24 inches so as not to obstruct a driver's view, are likely reasonable time, place and manner restrictions. Memorandum Re: CC File 050770/An Ordinance Relating to Roadside Memorials, Grant F. Lanley (October 12, 2005). In other words, if roadside memorials and messages could be erected grieving family members then other memorials and messages would also be allowed touting or promoting any idea, perspective or opinion on similar topics or on any topic.

In Milwaukee a substitute a resolution recognized that "citizens [may] grieve by placing a memorial, at the site of a fatal accident or occurrence within the boundaries of the public street, sidewalk or walkway" and noted that the Wisconsin Department of Transportation already implemented a policy that removed memorials that "interfere[d] with roadway safety, impact the free flow of traffic or fall into disrepair." The resolution also indicated the safety hazard resulting from the interference of "pedestrian and motor vehicle traffic through illegal loitering and littering within the boundaries of the public street, sidewalk or walkway." Finally the resolution directed the Milwaukee Department of Public Works to "develop a policy regarding roadside memorials" that would provide for the "removal of all obstructions that are in violation of current ordinances, including roadside memorials, no later than 30 days of the Department of Public Works being notified of their existence." The Milwaukee Common Council adopted the Resolution on March 23, 2006 and the Mayor Tom Barrett signed the ordinance in law seven days later. City of Milwaukee, Legislative Reference Bureau File Number: 051413.

City of Milwaukee Legislative Reference Bureau Memo entitled "Legislation relating to roadside memorials in various communities" and dated September 23, 2005, indicated that at that time cities in Massachusetts (Leominster: 90 day memorial or permanent traffic safety sign requested with 6 months), Illinois (Chicago: case-by case dismantling), California (Oakland: dismantling of roadside memorials with 24 hour) and Maryland (Poolesville). Other states allow some form of memorial: Colorado (removal of uniform sign after three years), New Jersey (removal of impromptu memorial after 10 days), Alaska (allows

impromptu memorials), Idaho (uniform gold stars and memorials limited by size and weight), and West Virginia (registration of permanent memorials). Legislation Relating to Roadside Memorials in Various Communities, City of Milwaukee Legislative Reference Bureau Memo (September 23, 2005).

Official Documents of State and Local Governments and Geographic Data

Considering the documents sourced from government sources and reproduced on the various social media or cycling websites issues relating to the copyright status emerge. Some countries legislate that such documentary record of the government at any level are part of the public domain. Such works are outside the subject matter of copyright and reside in the public domain.

[Vytautas Mizaras, 2 Copyright Throughout the World § 24:13 (Database updated November 2015) (§ 24:13. Works excluded from protection: "Pursuant to Article 5 of the Copyright Law, the subject matter of copyright protection does not include the following: (1) legal acts, official documents, and texts of administrative, legal, or regulative nature (decisions, rulings, regulations, norms, territorial planning, and other official documents), as well as their official translations; (2) official state symbols and insignia (flags, coat-of-arms, anthems, banknote designs, and other state symbols and insignia), the protection of which is regulated by other legal acts; (3) officially registered drafts of legal acts; (4) regular information reports on events; and (5) folklore."]

While works of the federal government are designated by statute to reside outside the subject matter of copyright ["Copyright protection under this title is not available for any work of the United States Government, but the United States Government is not precluded from receiving and holding copyrights transferred to it by assignment, bequest, or otherwise." 17 U.S.C.A. § 105.], it is the courts that have made pronouncements expanding the scope of the public domain into state court, legislative and regulatory documents. "Statutes are in the public domain and cannot be copyrighted." *Wheaton v. Peters,* 33 U.S. 591 (8 Peters) (1834); "It is well settled that judicial opinions and statutes are in the public domain and are not subject to copyright." *Veeck v. Southern Bldg. Code Congress Intern. Inc.*, 241 F.3d 398, 412 (5th Cir. 2001).

Public policy dictates that the documentary building blocks of law and government; its cases, statutes, regulations, etc. are in the public domain. The rationale for this is as follows: "two considerations influence whether a particular work may be properly deemed in the public domain: (1) whether the entity or individual who created the work needs an economic incentive to create or has a proprietary interest in creating the work and (2) whether the public needs notice of this particular work to have notice of the law." *County of Suffolk, New York v. First American Real Estate Solutions*, 261 F.3d 179, 194 (2d Cir. 2001). In the case study before us the crime data are not protected by copyright as purely factual content unless selected, coordinated or arranged in some creative manner as a compilation copyright. It could be argued that the remaining documentary record of city reports, planning memorandum, meeting minutes of governmental bodies are therefore in the public domain. Both criteria are satisfied. The governmental bodies creating the documents will continue to produce these with without any economic incentive to do so or not. Second, reports, planning documents, agendas and official minutes serve to notify the public of matters of concern.

General route and map information would not appear to fall into this cluster of works.

The determining factor would be whether "the existence and content of Suffolk County's maps are purely dictated by law, [then] it is likely that Suffolk County needed no additional incentive to create them. As we have indicated that Suffolk County is entitled to present evidence whether its tax maps are original." *County of Suffolk, New York v. First American Real Estate Solutions,* 261 F.3d 179, 188 (2d Cir. 2001). Compare discussion of geographic data from two other states. The Seventh Circuit in a case involving tax assessment data from Wisconsin assumed that building block of tax assessment mapping, the "address, owner's name, the age of the property, its assessed valuation, the number and type of rooms, and so forth… [about which] municipalities collect such data in order to assess the value of the properties for property-tax purposes" were a part of the public domain beyond the

protection of the copyright tlaw. *Assessment Technologies of WI, LLC. v. Wiredata, Inc.*, 350 F.3d 640, 642 (7th Cir. 2003): "The copyright case seeks to block WIREdata from obtaining noncopyrighted data. AT claims that the data can't be extracted without infringement of its copyright. The copyright is of a compilation format, and the general issue that the appeal presents is the right of the owner of such a copyright to prevent his customers (that is, the copyright licensees) from disclosing the compiled data even if the data are in the public domain." In another case it was assumed that similar mapping data was in the public domain: "For public safety reasons, it is critical that geospatial information such as the GIS Basemap stay out of the public domain… The actual location of the Hetch Hetchy water lines are generally known, but not provided in any detail for obvious reasons—to minimize the threat of terrorist attack on the water system… The exact location of Hetch Hetchy water lines is an integral part of the GIS Basemap and not easily segregable." *County of Santa Clara v. Superior Court*, 89 Cal.Rptr.3d 374, 394 (Cal. App. Dist. 6, 2009).

The route and map information in the current case study would face a similar fate as long as no creative elements were present. As one court observed: "Thus, Suffolk County may own a copyright under the Copyright Act. The question remains whether Suffolk County has sufficiently alleged that it possesses a valid copyright in its tax maps." *County of Suffolk, New York v. First American Real Estate Solutions*, 261 F.3d 179, 188 (2d Cir. 2001).
The question would be whether "the overall manner in which [the plaintiff] selected, coordinated, and arranged the expressive elements in its map, including color, to depict the map's factual content." *Streetwise Maps, Inc. v. Vandam, Inc.,* 159 F.3d 739, 748 (2d Cir. 1998). Without further examination of each map or geospatial dataset it would be difficult to generalize whether the compilation of such crosses the low level of creativity required in its selection, coordination and arrangement to qualify for copyright protection. The data itself remains in the public domain.

The Right of Privacy: Intrusion in a Public Place?

"Indiana recognizes four separate forms of the tort of invasion of privacy: (1) appropriation; (2) intrusion; (3) public disclosure of private facts; and (4) false light in the public eye. *Cullison v. Medley*, 570 N.E.2d 27, 31 (Ind. 1991). Whether there can be an "invasion" of privacy in a public place depends on the circumstances. For example, in ***Wilkins v. National Broadcasting Co., Inc.***, 84 Cal.Rptr.2d 329, 336-337 (Cal.App. 1999), the court concluded that there was no intrusion when two defendant reporters posing as potential investors met with company representatives and surreptitiously recorded the conversation. The reporters brought two others with them, and the representatives did not question the two guests' presence. The representatives, reporters, and guests sat at table close to other tables on a crowded restaurant patio. The table was not secluded and the representatives spoke freely in sales fever pitch, even when waiters came by. In other words the circumstances revealed that the subjects did not consider the conversations to be private, moreover admitting that they provided the same pitch to hundreds of other potential investors. The court concluded: "Pursuant to our review of the videotape and consideration of the admissions of Wilkins and Scott, we conclude that Wilkins and Scott had no objective expectation of privacy in their business lunch meeting." ***Wilkins v. National Broadcasting Co., Inc.***, 84 Cal.Rptr.2d 329, 336 (Cal. App. 1999). This is in contrast to an example where a women walks over an exhaust grate and a gust of air blows her skirt to reveal her underwear and someone takes a photograph. (See, Restatement (Second) of Torts, 1977, § 652B, illustration 7) However, recording intimacies that would in the normal course of observation be perceptible would not. For example, when two individuals are observed and recorded in the throes of a passionate kiss while seated on a bench located on a pathway in a public park or the perhaps less risqué but more famous example of a photograph of then Lady Diana holding one of her nursery charges against a sun-lit background so that the outline of her legs and underwear were viewable through fabric of her sheer dress. Recording the observations of the apparent would not appear to be the sort of unexpected or unreasonable invasion the tort contemplates. Some commentators however, see less distinction: "If an action, no matter how embarrassing, is recorded or photographed in public, there can be no intrusion to seclusion because there are usually no indicia of privacy

for anything that takes place in public." L.J. Kutten and Frederic M. Wilf, 4 Computer Software Protection-Liability-Law-Forms § 19.5 (2005) (Intrusion to Seclusion) (October 2016 Update), citing *Gautier v. Pro-Fooball, Inc.*, 107 N.E.2d 485 (N.Y. 1952). The resolution may depend on what measure the subject to protect their privacy: wrapping a towel around oneself while changing out of biking gear into street clothes but the subject is disrobed by an unexpected gust of wind that is captured in image or video.

It would also seem that based on this case law and discussion found in the Restatement (Second) of Torts even if a subset of the public were self-selected because of list, board, blog or chat subject matter or through password subscription, comments made by forum participants would still appear to be made in "public view" without any sense of seclusion. Given the nature of online forums if participants desire a conversation to be private, then person-to-person email, private chat or some other means of secure communication should be used. In the author's opinion a court would not conclude a comment lifted from a posting on a public or open forum such as a list, board, blog or chat is cloaked with that expectation nor would its publication in another venue be an intrusion.

The Right of Privacy: Appropriation

A second privacy right is appropriation. When another "appropriates to his own use or benefit the name or likeness of another" an invasion of privacy occurs. (Restatement (Second) of Torts, 1977, § 652C, "Appropriation of Name or Likeness") "Although the protection of his personal feelings against mental distress is an important factor leading to a recognition of the rule, the right created by it is in the nature of a property right." (Restatement (Second) of Torts, 1977, § 652C, comment a) The privacy right of appropriation, or to be more accurate to control the appropriation of one's name or likeness is related to the property right of publicity that is often associated with celebrities, public figures or other famous people. Black's Law Dictionary defines a right of publicity as "[t]he right to control the use of one's own name, picture, or likeness and to prevent another from using it for commercial benefit without one's consent."
Though a property right, the use need not be commercial or pecuniary. "It applies also when the defendant makes use of the plaintiff's name or likeness for his own purposes and benefit, even though the use is not a commercial one, and even though the benefit sought to be obtained is not a pecuniary one." (Restatement (Second) of Torts, 1977, § 652C, comment b) However, the advantage must come not from the mere use of the name, i.e., identifying the speaker of a particular comment, but from the adoption of the name or likeness for some benefit. In the present discussion the appropriation is made by the researcher for his or her personal gain: "the defendant must have appropriated to his own use or benefit the reputation, prestige, social or commercial standing, public interest or other values of the plaintiff's name or likeness." (Restatement (Second) of Torts, 1977, § 652C, comment c)

The difference is not found in the nature of person nor in the events that trigger the right of privacy appropriation versus a right of publicity, as the same event can trigger either. Rather the difference is found in the nature of the harm each victim experiences and the legal rights providing remedy. The appropriation that brings injury to the person, is subjective (though measureable as well in dollars a cents) whereas the right of publicity represents a loss of commercially exploitable opportunities in plaintiff's name, likeness or appearance. The former right is a right of personhood while the latter is a right based in property. "[T]he invasion of privacy by appropriation of name or likeness is a personal right, while the right of publicity more closely resembles a property right created to protect commercial value." *Rose v. Triple Crown Nutrition, Inc.*, 2007 WL 707348 (M.D. Pa. 2007). The cyclists whose images or likeness is appropriated may suffer either harm but more like a privacy appropriation claim would apply to individuals unfamiliar with the use of their image or persona in a commercial setting.

In online settings appropriation may also occur. But such claims are difficult to make in the ubiquitous nature of online settings: "The plaintiff, Beverly Stayart ("Stayart"), conducted

search engine queries with her own name as a search term, and she didn't like the results. Her queries produced links to ***pornographic websites, online pharmacies promoting sexual dysfunction drugs, and an adult-oriented online dating service***. Stayart alleges that the defendants, Yahoo! Inc. ("Yahoo!"), Overture Services, Inc. ("Overture"), and Various, Inc. ("Various"), knowingly and intentionally used her name on the internet without authorization. *Stayart v. Yahoo! Inc.*, 651 F.Supp.2d 873 (E.D. Wis. 2009), affirmed 623 F.3d 436 (7th Cir. 2010). The plaintiff was unsuccessful. All claims were dismissed. Such uses in social media sites might likewise be deemed incidental: "The fact that the defendant is engaged in the business of publication, for example of a newspaper, out of which he makes or seeks to make a profit, is not enough to make the incidental publication a commercial use of the name or likeness. Thus a newspaper, although it is not a philanthropic institution, does not become liable under the rule stated in this Section to every person whose name or likeness it publishes." (Restatement (Second) of Torts, 1977, § 652C, comment d) Likewise the mere mention of a subject's name or known pseudonym or likeness in a newspaper, newsletter, or social media website would not appear to be an "appropriation" of the subject's name for the authoring entity's "purpose and benefit" such that it would constitute an invasion of privacy.

Conclusion

The main ethical and legal issues identified in this paper will inform the design of the community archive and the utilization of technology during its construction. As ethical considerations are less clear than legal ones, the co-creation of the archive with community members will help guide ethical decision-making about data collection, analysis, and preservation. This will be especially important when dealing with personal data or content that is contributed by citizens.

Appendix

Definitions from 17 U.S.C. § 101:

"Audiovisual works" are works that consist of a series of related images which are intrinsically intended to be shown by the use of machines, or devices such as projectors, viewers, or electronic equipment, together with accompanying sounds, if any, regardless of the nature of the material objects, such as films or tapes, in which the works are embodied.

A "compilation" is a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship. The term "compilation" includes collective works.

" Literary works" are works, other than audiovisual works, expressed in words, numbers, or other verbal or numerical symbols or indicia, regardless of the nature of the material objects, such as books, periodicals, manuscripts, phonorecords, film, tapes, disks, or cards, in which they are embodied.

"Pictorial, graphic, and sculptural" works include two-dimensional and three-dimensional works of fine, graphic, and applied art, photographs, prints and art reproductions, maps, globes, charts, diagrams, models, and technical drawings, including architectural plans. Such works shall include works of artistic craftsmanship insofar as their form but not their mechanical or utilitarian aspects are concerned; the design of a useful article, as defined in this section, shall be considered a pictorial, graphic, or sculptural work only if, and only to the extent that, such design incorporates pictorial, graphic, or sculptural features that can be identified separately from, and are capable of existing independently of, the utilitarian aspects of the article.

- Copyright
- Licensing
- RTPM : Regulating Ghost Bikes
- Public Domain : State and Local Governments
- Privacy Rights

The recent improvements on circulation of research results at the Japan Atomic Energy Agency (JAEA)

Satomi Inagaki, Misa Hayakawa, Naomi Ebisawa, Mayuki Gonda, Takashi Nozawa, and Keizo Itabashi
Japan Atomic Energy Agency, JAEA, Japan

Abstract
The sharing of information about scientific research results on the Internet has developed with the current global advancement of open science, including archiving and disseminating scientific papers in institutional repositories, facilitating access to and use of research data etc. Accessibility to such large volumes of information on the Internet is a very important issue. Without solving the accessibility issue, those contents may remain grey literature. This paper introduces the case study of the Japan Atomic Energy Agency (JAEA) Library as an example of its efforts to improve the circulation of research results in terms of grey literature. JAEA has disseminated information of our research results via the Internet for over decade, but three main issues remain to be solved, aimed at improving the accessibility of grey literature in the open science era; (1) to ensure accessibility of our Internet contents, (2) to consider how our target users find our contents, (3) to improve the user interface of our contents. Finally, we consider enriching the contents of the JAEA Reports and accelerating the circulation of the JAEA R&D results by paying attention to the global trend of open science.

1 Introduction

The Japan Atomic Energy Agency (JAEA) is a comprehensive R&D institute dedicated to nuclear energy in Japan. The JAEA Library is one of the largest nuclear information centers in Japan. The JAEA Library provides technical information services to researchers and engineers in the nuclear science field, including to JAEA staff members. Additionally, we also disseminate information about R&D results achieved by JAEA staff members. As part of this activity, we publish the JAEA Reports (technical reports of JAEA) and the “JAEA R&D Review” (annual publication that introduces JAEA’s R&D results). This paper describes the case study of the JAEA Library as an example of its efforts to improve the circulation of research results in terms of grey literature.

2 The JAEA Library

2.1 The Japan Atomic Energy Agency (JAEA)

JAEA’s mission is to contribute to the welfare and prosperity of human society through nuclear science and technology. JAEA has focused on issues such as the response to the accident at the Fukushima Daiichi Nuclear Power Station of the Tokyo Electric Power Company Holdings, Incorporated (the Fukushima Nuclear Accident), which is our highest priority, research for safety improvement in the field of nuclear energy, R&D toward the establishment of nuclear fuel cycle technology, and the development of technology for the treatment and disposal of radioactive wastes. In addition, to support these R&D and create new technologies, we have implemented basic nuclear science and engineering research and human resources development. JAEA’s largest site is located in Tokai-mura (approximately 150km north of Tokyo), where the Central Library is located. (Figure 1)

Figure 1 the Central Library of JAEA

2.2 Features of the JAEA Library

The JAEA Library is involved in a wide range of activities. First of all, as the largest nuclear information center in Japan, the JAEA Library holds and provides 50,000 books, 2,000 academic journals, and 770,000 technical reports in the fields of nuclear science and technology.

The JAEA Library also plays the role of the INIS National Centre for Japan. The International Nuclear Information System (INIS) is the world's largest collections of published information relating the peaceful usage of nuclear sciences and technology, and is operated by International Atomic Energy Agency (IAEA). The JAEA Library prepares bibliographic records for nuclear-related literature published in Japan, and subsequently submits these to the INIS database.

After the Fukushima Nuclear Accident (March 11, 2011), we have distributed reference information related to the accident. The Fukushima Nuclear Accident Archive (FNAA) constitutes our most recent activity. In order to collect and preserve comprehensive information regarding the accident and to utilize it anytime in the future, FNAA provides reliable information on the Internet, for example concerning the Tokyo Electric Power Company Holdings, Incorporated and the government agencies of Japan. Additionally, FNAA provides information about oral presentations at meetings of academic societies in Japan about the Fukushima Nuclear Accident. (Figure 2)

Figure 2 the Fukushima Nuclear Accident Archive (FNAA)

2.3 Dissemination of the JAEA R&D Results

One of the important missions of the JAEA Library is the dissemination of the R&D results that are achieved by JAEA staff members.

JAEA staff members submit approximately 1,800 papers that contribute to scientific journals and give approximately 2,500 oral presentations in a year. JAEA staff members have to submit a form to their directors using a web-based system when they plan to publish the JAEA Reports, submit a paper to a journal, or make a presentation at conferences. This form consists of bibliographic data (title, abstract, title of journal, name of conference, etc.). The form's data from JAEA's R&D departments is stored in the database, which is managed by the library. To maintain the data quality, the library staff members check the data and correct errors, create authority files (author names. journal title, etc.), and control the bibliographic data. After this check by the library staff members, the bibliographic data about the publications and presentations are made available on the JAEA Originated Papers Searching System (JOPSS), following the publication of JAEA Reports and journals or oral presentations at conferences. (Figure 3)

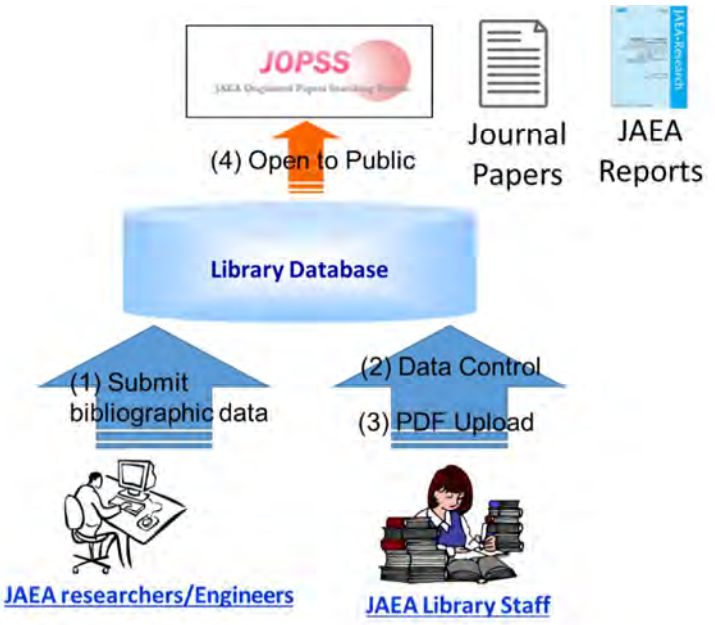


Figure 3 Management of the JAEA R&D results

2.4 The JAEA Reports

The JAEA Library publishes the JAEA Reports, which are technical reports of the JAEA and might be considered grey literature. (Figure 4) In the nuclear science field, the technical report is one of the important tools for exchanging R&D results. Nowadays, many technical reports are published on the websites of institutes or universities, which enables a better circulation of this kind of grey literature.

Figure 4 the JAEA Reports

The JAEA Reports are divided into seven types according to their report codes (or report prefixes). (Figure 5) The advantages of the JAEA Reports include the fact that (1) the R&D results can be described in detail because there is no limit to the number of pages, (2) the R&D results can be published while retaining their copyrights. Additionally, full-text PDF files of all JAEA Reports are provided on JOPSS, and (3) less time is required for publishing the JAEA Reports for traditional academic journals.

Report code	Definition
JAEA-Research	Research reports, Research review
JAEA-Technology	Technical reports
JAEA-Data/Code	Numerical data, Computer/Calculation code and Database
JAEA- Testing	Manuals and Experimental results
JAEA-Evaluation	Institutional evaluations, Research evaluations and Project completion report
JAEA-Review	Research review, Annual reports and Theses, etc.
JAEA-Conf	Conference Proceedings

Figure 5 Seven types of the JAEA Reports

2.5 The JAEA Originated Papers Searching System (JOPSS)

JAEA staff's R&D bibliographic information (academic papers and oral presentation) are available on JOPSS. Moreover, JOPSS provides bibliographic information and the full-text PDF files of the JAEA Reports. Approximately 22,000 full texts from the JAEA Reports are downloadable as of November 2016. (Figure 6)

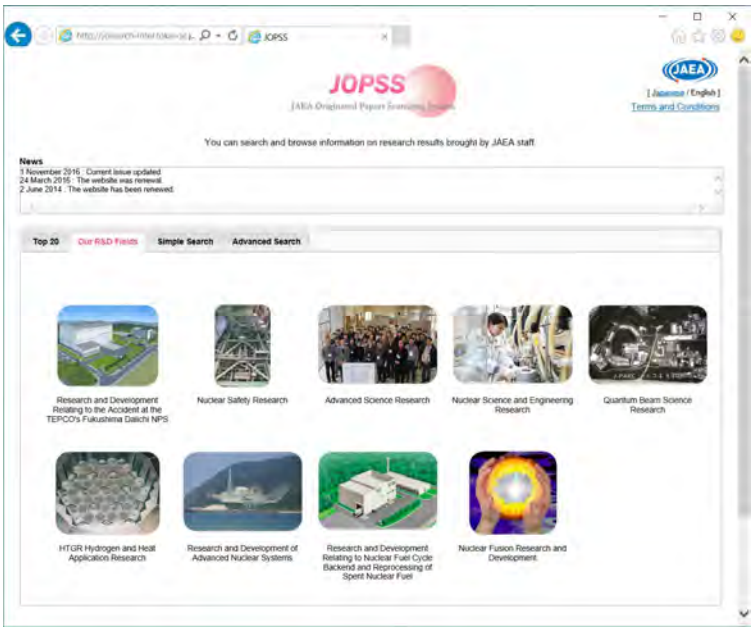


Figure 6 the JAEA Originated Paper Searching System (JOPSS)

JOPSS was launched in 2006 and is open to the public through the Internet. JAEA's R&D bibliographic information and the full texts of the JAEA Reports are available for everyone at any time. However, this accessibility is not very common and the JAEA Reports were still considered grey literature. Subsequently JOPSS was improved by means of several modifications with the aim of disseminating JAEA's research results more efficiently.

The following three main problems need to be solved if we aim to improve the accessibility of grey literature in the open science era:

(1) To ensure accessibility of the JAEA Reports
(2) To consider how our target users find our contents
(3) To improve the user interface of our contents

3 Modifications of JOPSS

3.1 To ensure accessibility of the JAEA Reports

To ensure the accessibility of the JAEA Reports, we have assigned a digital object identifier (DOI) to the JAEA Reports from 2014 onwards. (Figure 7) DOI is a persistent identifier (PID) that identifies electronic objects and provides the current location where these objects can be found. Because DOI does not change, it ensures a permanent access to the JAEA Reports on the Internet. DOI is used as reference URI (Uniform Resource Identifier). By using DOI, other databases can refer users to the JAEA Reports almost permanently. As a result, the JAEA Reports' circulation will improve.

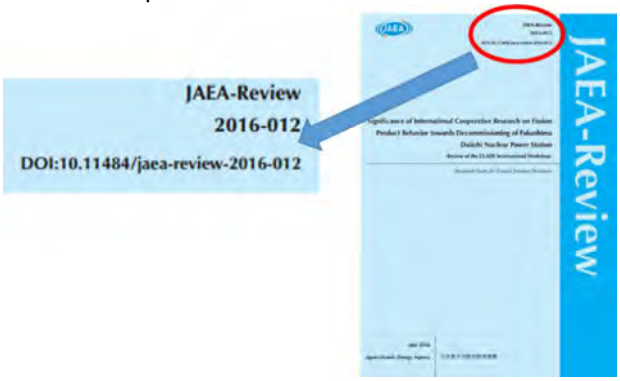


Figure 7 DOI of the JAEA Reports

3.2 To consider how our target users find our contents

To consider how our target users find our contents, we have added a direct access route to the metadata of the JOPSS contents. It is important that users can easily and correctly find what they want. Our target users are researchers and/or support staff members for researchers. They tend to use general search engines (e.g., Google) and major bibliographic databases to find information for their research activities.

However, the former JOPSS provided search results only in lists, without any independent detailed pages. Accordingly, search engines could not crawl all metadata of the contents. Therefore, we assigned a unique URL to each detailed page in JOPSS.

Additionally, JOPSS has started to connect with other major bibliographic databases. The Japanese Institutional Repositories Online (JAIRO) provides a cross searching service of institutional repositories in Japan. The contents of JOPSS are automatically harvested by JAIRO using a Web-API.

Subsequently, we provide direct access to the contents of JOPSS from search engines or other bibliographic databases. (Figure 8)

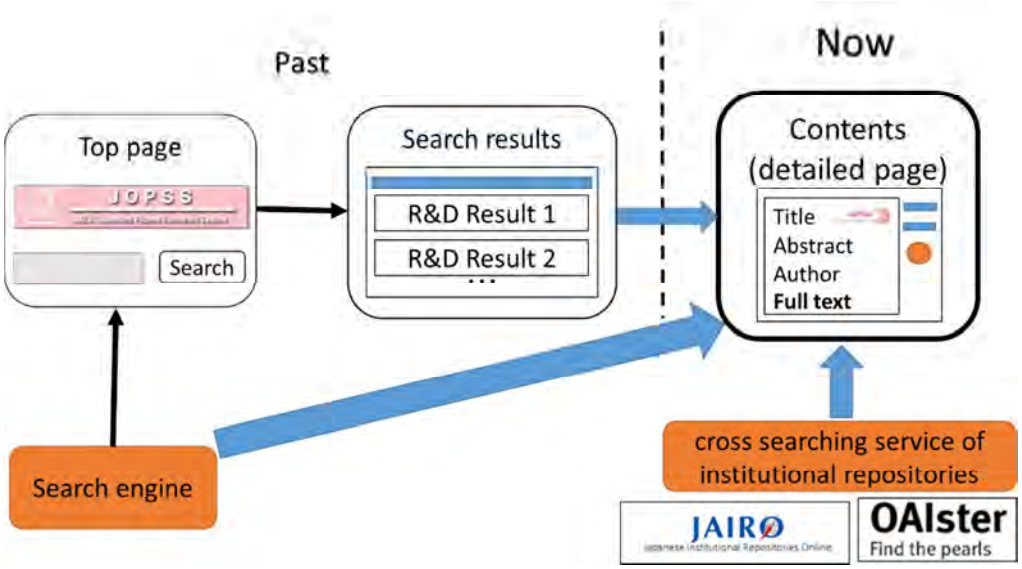


Figure 8 Direct access from search engines and cross searching services of institutional repositories

3.3 To improve user interface of our contents

To improve the user interface of our contents, we added the following functions. When using the former JOPSS, users needed to enter keywords to search the contents. This meant that they could not search intuitively. We now provide two new tabs on the top page of JOPSS, “Our R&D Fields” and “Top 20”, to indicate expectations concerning the contents. The “Our R&D Fields” tab contains photos or illustrations of our main 9 research fields. Each changes every few seconds and links to the list of contents (including bibliographic information of the research results) in the research field that is indicated by the photos or illustrations. (Figure 9) “Top 20” tab provides lists of the top 20 contents that were most frequently accessed and downloaded during the past 30 days (Figure 10), and these lists link to detailed pages about the contents. As a result, users can access such detailed pages directly from the top page with one click and without entering any keywords.

Figure 9 “Our R&D” tab

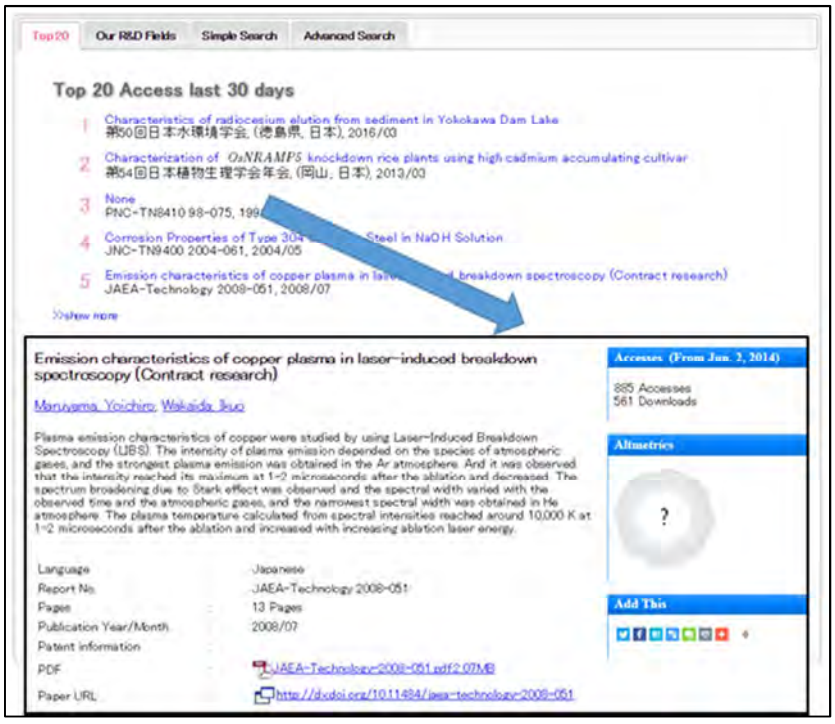


Figure 10 “Top 20” tab

To provide users with effective information, we added functions that show the contents’ impact. There are two types of effective information, namely, article metrics and related information about the research results. The times of access and downloads, the Web of Science’s Times Cited count, and the Altmetrics badge provide information about article metrics. (Figure 11) The Web of Science is an academic research information database that provides a citation index of the world’s most important research covering scholarly journals, books, proceedings, published data sets and patents. “Times Cited” is a one of the Web of Science’s features. Altmetrics is the creation and study of new metrics, based on the Social Web, for analyzing and informing scholarship. The DOI of each journal paper plays an important role in showing the Web of Science’s “Times Cited” count and embedding the Altmetrics badge. We use the DOI as identifier of the article to get this information from other database/Web services. Library staff members maintain the bibliographic data to enable the exact collection of the DOI. This is an example of how the DOI is used effectively.

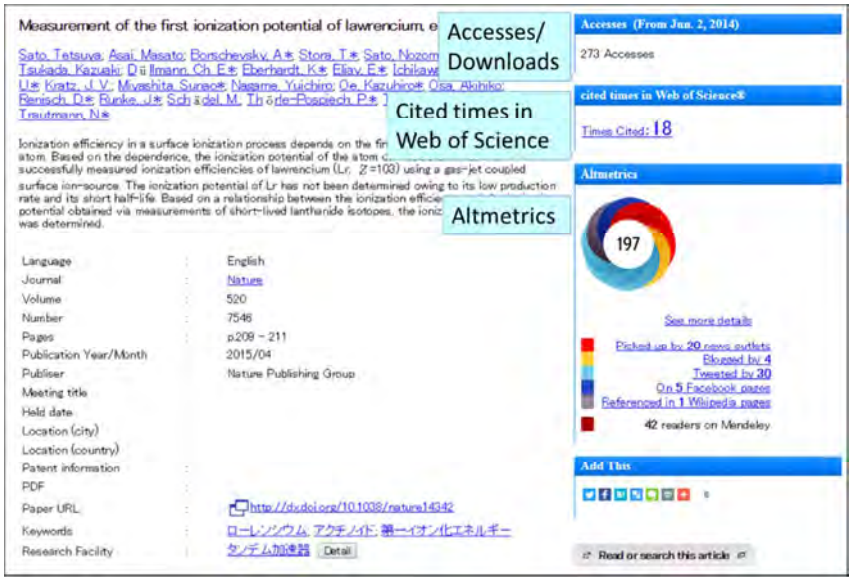


Figure 11 article metrics on the JOPSS

In addition, for related information about the research results, we provide links to the web page of the facilities that were used for the research as well as links to related patent information. (Figure 12)

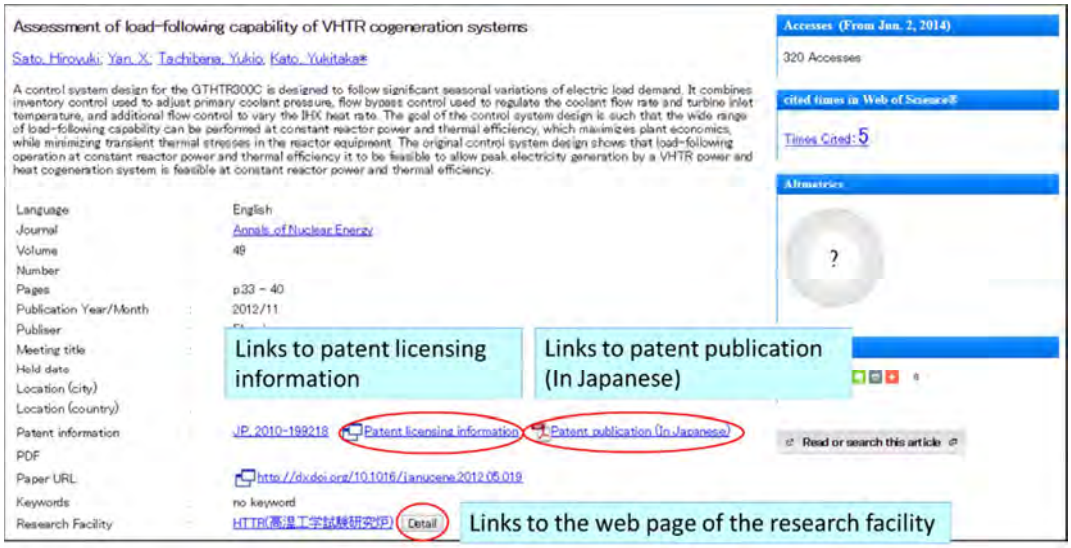


Figure 12 related information on the JOPSS

4 Future Plans

The JAEA Reports have been made accessible on the Internet through JOPSS. As mentioned in chapter 3, we have improved this system to ensure its accessibility and to improve the user interface.

To improve accessibility, we assigned a DOI to the each JAEA Report to provide users with permanent access and a reference URI. Therefore, researchers who know that the JAEA Reports exist can access the full- texts through a simple search.

However, the JAEA Reports are still “grey” for some potential users, such as researchers in fields other than nuclear science, industrial fields and foreign users. The amount of published information and data on the Internet increases rapidly nowadays. As a result, information that cannot reach users equals non-existence on the Internet.

The JAEA Reports might be useful not only for researchers in the field of nuclear science, but also for potential users. Therefore, we consider promoting the JAEA R&D results for these potential users and leading to use the JAEA Reports as reference.

Additionally, the correspondence of “Open Science” needs to be examined. The movements of “Open Science” have been a point of focus all across the world. In Japan particularly,

"Open Science" is a concept that encompasses open access by turning research data into open data. Therefore, opening research data to the public from JAEA is one of our future tasks.

At JAEA, we have published some research data as the JAEA Reports. However, there are some issues concerning how to disseminate them in terms of data sharing and archiving in open science. The research data in the JAEA Reports do not have a metadata format and are therefore not searchable. As a result, these data are still "grey" for potential users.

For the first step of research data dissemination, we will make guidelines that indicate how to process research data and to write descriptions for JAEA research data publication.

As described above, we plan to enrich the contents of the JAEA Reports and accelerate the circulation of the JAEA R&D results.

5 Conclusion

Scientific research results on the Internet play the important role of sharing information, such as archiving and disseminating scientific papers in institutional repositories and facilitating access to research data in open science. It may be true that it is easy for users to reach to research results because the Internet is usually open to the public. However, information on the Internet does not have permanent accessibility. Additionally, it is difficult for searchers to get access to scientific research results in the form of technical reports in fields other than their own. Without solving this accessibility issue, "grey" literature will remain on the Internet.

This paper presented our attempt to improve the circulation of the JAEA R&D results in terms of grey literature. We continue to disseminate research results by paying attention to the global trend of open science.

We hope that our work will contribute to leverage diversity on grey literature in open science.

References

Clarivate Analytics. Web of Science. http://ipscience.thomsonreuters.com/product/web-of-science/

Gonda, M., Kunii, K., Nakajima, H., Ikeda, K., Itabashi, K. 2011. Circulation Improvement of Articles in Journals written in Non-English Languages; A Special Journal Titles Translation List of Journals written in Japanese for the International Bibliographical Database. Proceedings of 12th International Conference on Grey Literature (GL-12), 113-117.

Hayakawa, M., Nagaya, S., Ebisawa, N., Inagaki, S., Nozawa, T., Itabashi, K. 2015. Improving the Metadata of Bibliographic Information to Enhance its Access. Proceedings of 4th International Conference of Asian Special Libraries (ICoASL 2015), 8p.

Hayakawa, M., Kumazaki, Y., Nakajima, H., Yonezawa, M. 2016.Activities for preserving information resources of Fukushima Daiichi Nuclear Power Station accident; Constructing, operating the archive and its future prospect. [In Japanese], Journal of the Atomic Energy Society of Japan: ATOMOΣ, 58(8), 509-513.

Ikeda, K., Gonda, M., Nozawa, T., Itabashi, K., Ebisawa, N. 2009. Dissemination of Information on the Results of R&D through the JAEA Library. Proceedings of 11th International Conference on Grey Literature (GL-11), 111-117.

Ikeda, K., Ohshima, T., Gonda, M., Hayakawa, M., Mineo, Y., Yonezawa, M., Itabashi, K. 2014. Contribution to the improvement of dissemination of grey literature - JAEA Library's efforts for collecting, organizing and disseminating information on nuclear accidents. Proceedings of 15th International Conference on Grey Literature (GL-15), 41-47.

Ikeda, K., Gonda, M., Nagaya, S., Hayakawa, M., Mineo, Y., Kunii, K., Yonezawa, M., Itabashi, K. 2015. Enhancement of the functions of the Japan Atomic Energy Agency Library's Fukushima Nuclear Accident Archive using a novel data flagging system that improves the utilization of numerical data on the Internet. Proceedings of 16th International Conference on Grey Literature (GL-16), 139-145.

Sexuality Leveraged Through Diversity: Recognizing LGBT+ Communities

Anthony Lin, Irvine Valley College,
Julia Gelfand, University of California, Irvine, United States

Abstract
Gender studies has embraced a wide band of issues related to sexual identity, belonging, community, self-perception, transitional status, perceptions by the media and the changing political landscape of acceptance of all queer members. Recent political events such as the AIDS crisis, ability to self-define one's sexual orientation, marry and establish domestic partnerships, create families through adoption, increase in workplace accommodations, and expansion of the military, are but some of the milestones that have redefined opportunities for a significant population. The expanding gay, lesbian, bisexual, transgender, intersex, asexual communities and their allies have documented more inclusive categories of acceptance in and by society at large. This path had many diversions in the literature, media and true life experiences, leading to a more open and visible lifestyle. Elements of greyness pepper the changes as the body politik emerges into rightful social acceptance. Social and religious undertones define and defy the status quo that protects all members with legal rights and opportunities. The spectrum of grey parallels the colorful rainbow symbolizing sexuality, life, healing, sunlight, nature, magic/art, serenity/harmony and spirit that has come to depict the range of inclusivity that transforms society. Political and religious sectors globally influence the social injustices that members of this diverse community experience. Social media, news, film, literature, and the arts each chronicles life and its nuances by promoting more discourse and sharing about experiences that were once taboo public issues. The campaign for public awareness for the LGBT+ community has changed social norms redefining gender in a more plural, less polarizing context. Grey literature was once the alternative source for earlier study of these issues which has transitioned to an increasingly open and interdisciplinary range of content. This paper explores how society leverages change by defining new social practices and etiquette that translates into advocacy, civility, acceptance and a new extended definition of community, family and self. The potential for more sophisticated social media to leverage diversity remains ongoing. The celebration and embrace of these LGBT+ communities reduces the grey tinge while leveraging more candor in treating sexuality. The role of social media, once considered very grey is also mainstreaming into more common and established venues for capturing, publishing and sharing of personal and community experiences. Gender and sexuality through storytelling and other avenues of social media contributes to new understandings of complex human nature with greater openness.

Introduction
Exploring sexuality and diversity has a long history, with speculation going back to Biblical times. The focus of this research is how society came to recognize the power and strength of LGBT+ individuals and communities and the influence and challenges it brings to our current socio-economic and political landscapes via different communication strategies, flirting as grey literature yet at times illustrating how mainstream it has become. What we know best is the North American, specifically American experience of how the gay and lesbian movements developed much in response to the grand social changes initiated by World War II and the aftermath with redefining the traditional "American Family." Laws, regulations and policies dictated social norms as the sense of gay communities evolved with the early focus on gays and lesbians but soon expanded to include "bisexual, transgender, queer, questioning, and intersex organizing."[i]
Sexual identity is often divided up by how the history of gay and lesbian movements evolved giving sexual orientation a home in many camps. This includes the times when sexuality was not necessarily benign but silent, just not common dinner table conversation and was described as "in the closet" if one did not subscribe to traditional heterosexual practices and identity.

The Rainbow has become the mascot and symbol of gay pride and rights. Bright colors were used to self-identify and share information in the early days. In 1978 the first gay pride flag designed by San Francisco artist, Gilbert Baker combined eight colors each reflecting its own symbolism and remaining the iconic symbol of the LGBT community since then:

- Pink for sex
- Red for life
- Orange for healing
- Yellow for sunlight
- Green for nature
- Blue for art
- Indigo for harmony
- Violet for the human spirit

Eventually due to availability and restricting the rainbow to six colors, pink and indigo were dropped. The rainbow resonance also has symbolism from the Bible and other forms of iconography.[ii]

Due to several events, the coming out era transformed public opinion. Here in New York in the summer of 1969 Stonewall became what is recognized as the single most important event leading to the gay liberation movement and the modern fight for LGBT+ rights in the United States. Riots and uprisings defined that summer and for the first time, creating safe places for the gay community outside of traditional bars where relationships between local police and the gay community were tested. Within a year, two gay activist organizations, and three gay newspapers were established followed by gay rights organizations and later that year the first Gay Pride marches were organized in major cities to commemorate the anniversary of Stonewall and today that is a common celebration in late June of how far gay rights has come as they are common events in many cities, now worldwide.

Evolution of Terminology for Sexuality: Identity & Changing Sexual Roles

The model of gender and sexual identification that most of us are familiar with is the duality of gender (female and male) and the heterosexual or homosexual orientation. The term for this model of heterosexual identification is called, cisgender, in other words, people who are not transgender and identify with their biological sex, the majority of the human population. Transgender identification encompasses a vast range of biological, sexual, and societal expressions. Biologically, one can be born with either female or male sexual organs, or as intersexuals where duality of sexual organs is exhibited and sex characteristics that do not fit typical binary notions of bodies designated “male” or female.” Sexual orientation can include being attracted to the opposite sex (heterosexual), the same sex (homosexual), both sexes (bisexual), or no one particular sex (asexual). In terms of societal expression, one can choose to express themselves as female or male in clothing and mannerisms but be from the opposite biological gender or sexual orientation. Keeping it simple, our definition for transgender refers to people whose sense of their gender is different from their sex at birth. Contemporary examples of gender-neutral expressions are increasingly evident in daily life as challenges define privacy issues for public facilities and lifestyle choices.

This leads to difficulties in properly classifying a person’s identity and English language at this time, lacks proper pronouns to encompass the transgender identities beyond the commonly held fe/male gender model. Icons now reflect gender neutral options and common exchanges via texting has all but eliminated pronouns. We can see from tables included in the “Guide to Transgender Terms,”[iii] how language and vocabulary dictate social understanding. Due to the complex nature of transgender identification, this tends to lend itself in a grey area of classification as something not easily definable and in a dynamic state. This tends to lend itself to layers of intersectionalities[iv] not only in gender, sexual orientation, and social expression but also societal roles. For example, one can be biologically female, bisexual, express themselves as a male, and take on multiple roles of parent, sibling and profession. These intersections add to an even more complex set of interactions on top of an already complicated and often times not easily definable identity.

These multiple roles lead to new opportunities for LGBT+ persons to share common social and legal experiences of maturation through the lifespan. Today in many parts of the world, the LGBT+ population participates in marriage, military service, parenthood, durable power of health-care, beneficiary survivorship, are well represented in professional sports (hockey, tennis, American football, boxing, etc) and other roles related to the lifespan.

The issue of sexuality is complicated, very grey and constantly evolving as society evolves. The term LGBTQ+ evolved from GLBT in the late 1980s.[v] The letters "G" and "L" were switched because it was noted that the term gay, although it encompasses both men and women, next to the letter "L," the "G" really refers to "gay men." In some circles, it was thought that men are overrepresented and were overly dominant, so "L" was placed first, followed by the "G." Other groups who did not fall under, lesbian, gay, bisexual, or transgender orientations noted that they were not represented in the acronym, so more letters were added at the end. The "+" represents the inclusion of all groups for the sake of simplicity and inclusiveness.

Guide to Transgender Terms

Term	Definition
Cisgender	People who are not transgender.
Cross-dresser/ transvestite	People who wear clothes associated with the opposite sex but who do not necessarily identify with the opposite gender. (The older term "transvestite" is now widely viewed as pejorative.) Many male cross-dressers prefer female sex partners.
Gender dysphoria	Condition in which a person's birth gender and gender identity differ.
Gender identity	Internal, deeply held sense of one's gender.
Gender expression	Outward portrayal of a person's gender through name, pronouns, clothing, behavior, voice or body characteristics.
Gender nonconforming	People — not necessarily transgender — whose gender expression differs from conventional femininity or masculinity.
Sex-reassignment (or gender-confirmation) surgery	Surgery to change one's genital organs to those of the opposite sex; not undertaken by all transgender people.
Genderqueer	People who do not identify as either male or female.
Sex	Bodily characteristics — including chromosomes, hormones and reproductive organs — that define a person as male or female.
Sexual orientation	A person's enduring physical, romantic or emotional attraction to other people.
Transgender	People whose gender identity and/or gender expression differs from conventions associated with their sex at birth. Transgender people can be straight, homosexual or bisexual.
Transition	Period in which transgender people change their name and appearance.
Pronoun usage	Some transgender people prefer to be referred to by gender-neutral pronouns such as "they" and honorifics such as "Mx."

Sources: "GLAAD Media Reference Guide — Transgender Issues," undated, GLAAD, http://tinyurl.com/6u7k67c; "Gender Dysphoria," Diagnostic and Statistical Manual of Mental Disorders (DSM-5), American Psychiatric Association, 2013, http://tinyurl.com/kxaf3m6

Terminology has changed from the references dating from the 19th century where they were more sexology focused, using homosexual until the early-mid 20th century when gay became the accepted term reflecting the Stonewall generation and continues today. Contributions from the feminist movement, including vocabulary such as "queer" and by the 21st century, the + (plus) to illustrate more inclusive identities. Additional descriptors apply to those on the fringe as well and include Questioning, Intersex, Asexual, Allies and Pansexual.[vi]

The number of orientations span the full range of conventions and even the undefinable sexual orientations as seen below, to which are occasionally added new vocabulary and labels, extending to new meanings of the plus.[vii]

- Androgynous | Identifying and/or presenting as neither distinguishably masculine nor feminine.
- Asexual | The lack of a sexual attraction or desire for other people.
- Bisexual | A person emotionally, romantically or sexually attracted to more than one sex, gender or gender identity though not necessarily simultaneously, in the same way or to the same degree.
- Cisgender | A term used to describe a person whose gender identity aligns with those typically associated with the sex assigned to them at birth.
- Gay | A person who is emotionally, romantically or sexually attracted to members of the same gender.
- Gender dysphoria | Clinicallysignificant distress caused when a person's assigned birth gender is not the same as the one with which they identify. According to the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM), the term - which replaces Gender Identity Disorder - "is intended to better characterize the experiences of affected children, adolescents, and adults."
- Gender-fluid | According to the Oxford English Dictionary, a person who does not identify with a single fixed gender; of or relating to a person having or expressing a fluid or unfixed gender identity.
- Gender identity | One's innermost concept of self as male, female, a blend of both or neither – how individuals perceive themselves and what they call themselves. One's gender identity can be the same or different from their sex assigned at birth.

- Gender non-conforming | A broad term referring to people who do not behave in a way that conforms to the traditional expectations of their gender, or whose gender expression does not fit neatly into a category.
- Genderqueer | Genderqueer people typically reject notions of static categories of gender and embrace a fluidity of gender identity and often, though not always, sexual orientation. People who identify as "genderqueer" may see themselves as being both male and female, neither male nor female or as falling completely outside these categories.
- Gender transition | The process by which some people strive to more closely align their internal knowledge of gender with its outward appearance. Some people socially transition, whereby they might begin dressing, using names and pronouns and/or be socially recognized as another gender. Others undergo physical transitions in which they modify their bodies through medical interventions.
- Lesbian | A woman who is emotionally, romantically or sexually attracted to other women.
- LGBTQ | An acronym for “lesbian, gay, bisexual, transgender and queer.”
- Queer | A term people often use to express fluid identities and orientations. Often used interchangeably with "LGBTQ."
- Sexual orientation | An inherent or immutable enduring emotional, romantic or sexual attraction to other people.
- Transgender | Denoting or relating to a person whose sense of personal identity and gender does not correspond with their birth sex Therefore, transgender people may identify as straight, gay, lesbian, bisexual.
- Transphobia | The fear and hatred of, or discomfort with, transgender people.
- Transsexual |Persons who emotionally and psychologically feel that they belong to the opposite sex.

Changing hues and stereotypes: Depictions in the media and societal acceptance

The traditional depiction of LGBT characters in the popular media focuses on a two-dimensional representation of lesbians and gay males. On one side, the stereotypical gay male portrayed in the literature and media is usually characterized as Caucasian, effeminate, hedonistic, loves fashion, shopping, shallow personality, promiscuous, has a majority of female (straight) friends within his inner circle. On the other side is the stereotypical lesbian who is also White, but masculine oriented, enjoys the great outdoors, wears masculine clothing, is in a dysfunctional monogamous relationship, and has a majority of female (lesbian) friends within her inner circle. Stereotypes in any minority community contain a grain of truth and there are plenty of individuals who may have all of these traits. There are also plenty of exceptions to the rule that the media has only recently begun to acknowledge and portray. The attraction of lesbian couples often illustrate how each partner assumes traditional roles, one more male-dominant and one more feminine. However, among gay couples, the gender roles may be less visible and less obvious. It is possible that the media intends to play off of these stereotypes to draw in viewers and demonstrate extreme behaviors.

Sexual minorities often lie on the fringes of dominant society and US based television programming such as RuPaul’s *Drag Race*, *Will and Grace*, the popular sitcom, *Modern Family* that perhaps more than any one program swayed and influenced public opinion about gay marriage[viii], plus other avenues that provide a source of entertainment, spice, and exoticism that LGBT+ characters often portray. This leads to a “double edged sword” as the increase in LGBT+ programming leads to a greater familiarity of LGBT+ culture, however slanted it can be. At the same time, profits and entertainment value are the driving force for successful TV programming over actual, balanced portrayals of LGBT individuals and their stories or circumstances.

The recent release of Gregory Woods’ acclaimed work, *Homintern: How Gay Culture Liberated the Modern World* explores the role of the homintern as the international presence of lesbians and gay men in modern life, particularly in the creative fields of literature, cinema, music, dance and the arts. Predictably, it can be argued how it is likely one of the major creative forces in the cultural development of the past century.[ix] Woods

illustrates how certain cities including New York's Harlem neighborhood just up the west side of Manhattan in the 1910s was a haven for the presence of lesbians and gay men in the artistic avant-garde. "By loving as they chose, they reminded the world that love is worth taking risks for. By living as they chose, they initiated a process of liberation by sleight of hand."[x]

According to the *2016-2017 GLADD Where we are on TV Report*, overall there has been an increase in the number of LGBT+ characters on broadcast TV in the US. However, more female lesbian and bisexual characters were "killed" this particular TV season indicating a greater need for representation in the future. However, there was an increase shown in trans-characters that were played by actresses that are themselves transgender.[xi]

Summary of 2016 LGBT Characters

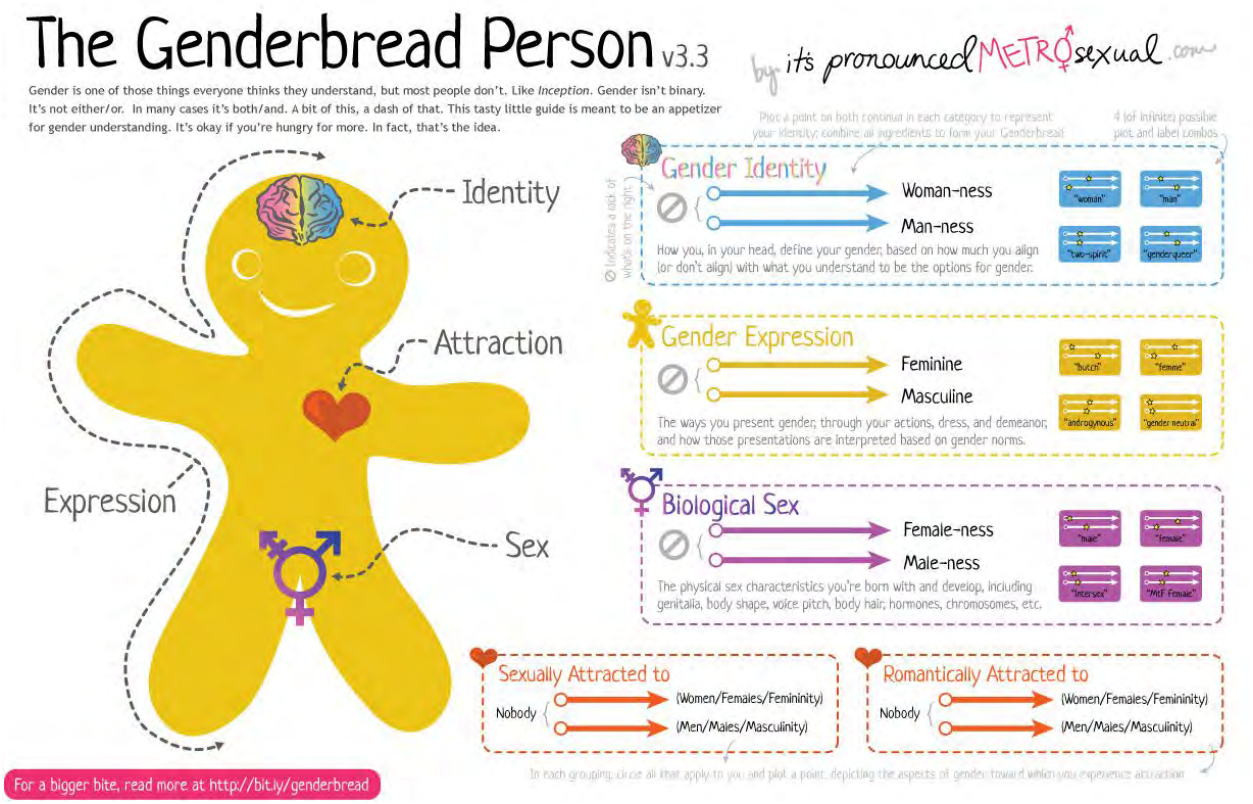
	Broadcast TV		**Cable**		**Streaming Media**	
	Comparison to 2005	**Number of Characters %**	**Number of Characters**	**Percentage**	**Number of Characters**	**Percentage**
Lesbian	3 /19%	12 /17%	29	20%	28	43%
Gay	12 / 75%	35/ 49%	65	46%	15	23%
Bisexual Female	1 / 6%	16 /23%	35	25%	13	20%
Bisexual Male	0	5 / 7%	10	7%	3	6%
Transgender	0	3 /4%				
Transgender Female			2	1%	7	11%
Transgender Male			4	3%		

Source: 2016 Where We are on TV, GLADD; Source: 2005 Where We are on TV, GLADD

If we compare the data above to the start of the *GLADD Where We are on TV Report* from a decade earlier (the 2005-06 statistics),[xii] we can see a steady increase of LGBT+ representation over the past 10 years including the inclusion of more trans-characters in TV, cable, and streaming media formats due to more acquired scripts showcasing such roles.

With the gradual increased visibility of LGBT active life in mainstream US culture, it is probable that we will continue to see increases in all lesbian, gay, bisexual, and trans portrayals in media overall. At some point in the future, it is hopeful that LGBT+ culture may become less controversial and further assimilated into society as a whole and will follow a similar path that earlier Civil and Human Rights movements took and eventually will lead to the integration and inclusion of LGBT+ culture as part of conventional mainstream society.

Due to the constantly growing list of identities and sexual orientation terms, Sam Killerman, created the "Genderbread Person." "The Genderbread Person is an infographic that breaks down gender identity, gender expression, biological sex, and sexual orientation into an easy to understand visual." As we can see from the graphic, sexuality is divided into four main categories: identity, sexual attraction, biological sex, and expression. There is a continuum of "man-ness" and "woman-ness", female/male assigned to each category. In addition, to account for asexual individuals, there is also a "nobody" in terms of sexual and romantic attraction.



Source: Genderbread Person v3

Although it is not mentioned, there is also an "undefinable" category for individuals who do not care to be "labeled" in any particular way, even though it could be argued that "undefinable" is also a category as well. There are no clear cut explanations when discussing sexuality as it is a very grey area that reflects the complex and evolving nature of humanity.

Developing Community: Different Scenarios

Animals and people seem to have the need to be around their own. This coupling and community forming nature creates a sense of security, family and belonging. The LGBT+ communities have grown and expanded into mature self-sufficient, self-governed or self-led diverse environments. Initially not always feeling welcome in many social settings, new membership driven associations and fellowships evolved over the last half century, especially in faith-based settings when traditional synods were reluctant to offer a welcoming atmosphere for LGBT+ parishioners with full participatory rights.

More liberal Jewish communities were among the first to form new congregations serving these communities and many of those have morphed into mainstream synagogues serving LGBT+ and heterosexual members. Beth Chayim Chadishim, opened in 1972 in Los Angeles as the world's first Synagogue serving the gay community and was at the forefront of the HIV crisis ministering to a growing need. It began celebrating the Transgender Day of Remembrance in November 1978 to memorialize those who were killed due to anti-transgender hatred or prejudice and statistics indicate that at least one such casualty is noted each month.[xiii] With the exception of the orthodox denomination, ordination of LGBT+ rabbinical clergy has been well established worldwide since the latter half of the last century. When describing the transformations of Jewish and LGBT+ life, "progressive yet misunderstood, often isolated, people"[xiv] characterizes the orthodox experience but suggests that times have changed in other persuasions of Judaism.

Within the Christian Community, the Metropolitan Community Church, also known as the Universal Fellowship of Metropolitan Community Churches has been at the forefront being inclusive of and providing extensive outreach to the LGBT+ communities. The Fellowship has Official Observer status within the World Council of Churches but has been denied membership in the US National Council of Churches due to its outreach efforts. Still, many local MCC congregations are members of local ecumenical partnerships around the world and MCC currently belongs to several statewide councils of churches in the United States.

The Unitarian Church has also been actively open to members' diversity. The Episcopal Church has consecrated gay/lesbian bishops and in 2015 began performing same-sex marriages.

The evangelical fundamentalists have traditionally been less accepting of LGBT outreach and inclusion due to strict fervent interpretations of the New Testament. However the Evangelical Network (TEN) is an association of LGBT+ affirming evangelical ministries and individuals and held its national conference in Irvine, California earlier this summer. TEN's mission is focused on the LGBT+ community and its allies to:

- Provide a safe place for LGBT+ people to discover and rediscover Christianity
- Bring together people with a shared faith to worship, evangelize and disciple
- Give people from all different backgrounds an opportunity to be educated about Homosexuality and Christianity[xv]
- Afford ministry and growth opportunities to people in the LGBT+ community
- Hold an annual conference open to all that is designed for networking, fellowship, education and worship
- Offer a safe place for LGBT+ people and the evangelical church community to dialogue
- Be a voice in the public and media that supports the LGBT+ community and its struggles for equality, and to speak out against negative stereotypes.

The Catholic Church remains steadfast in dealing with marriage, family, and sexuality, and doesn't move an inch on same-sex relationships or other LGBT+ issues. Pope Francis's release of the *Joy of Love* document earlier this year reaffirms the Roman Catholic Church's opposition to same-sex relationships, saying these can never be considered the equal of heterosexual marriage. Locally, many Catholic parishes have reached out to Catholic members but equality in the Church remains impossible at this time.

Muslims for Progressive Values is a faith-based, grassroots, human rights organization that embodies and advocates for the traditional Qur'anic values of social justice and equality for all, in the 21st Century. Its principles state: "We endorse the human and civil rights of lesbian, gay, bisexual, transgender, queer, and intersex (LGBTQI) individuals. We affirm our commitment to ending discrimination based on sexual orientation and gender identity and we support full equality and inclusion of all individuals, regardless of sexual orientation or gender identity, in society and in the Muslim community."[xvi]

Religion and spirituality have come a long way in admitting that its members in every facet of religious life is reflective of the LGBT+ population and its future is inherent in how to best minister to and support its members. A more open progressive atmosphere has been created but reasons of doubt about its future always remain under the clouds of a doubting self-righteous and conservative political base.

Census Self-Identification – Issues with Accurate Data

Today, we see how every walk of life, ranging from politicians, celebrities, athletes, law enforcement, civil servants, business people, academics, clerics who have embraced their own sexual orientation and can find that self-identifying is a powerful disclosure. New laws and perceptions of harassment argue for personal civil rights and will legally challenge situations of discrimination based on sexual identification.

The Williams Institute at the UCLA Law School issued in 2011 a demographic snapshot that an estimated 3.5% of adults in the United States identify as lesbian, gay, or bisexual and an estimated 0.3% of adults are transgender. This transcribes roughly to 9 million LGBT+ Americans and population analogies match it to the equivalent of the state of New Jersey.[xvii]

Other findings note that among adults who identify as lesbian, gay or bisexual, bisexuals comprise a slight majority (1.8% to 1.7% who identify as lesbian or gay); women are substantially more likely than men to identify as bisexual, estimates of those who report any lifetime same-sex sexual behavior and any same-sex sexual attraction are higher than estimates of those who identify as lesbian, gay or bisexual. There are also nearly 700,000 transgender individuals in the U.S. An estimated 19 million Americans (8.2%) report that they have engaged in same-sex sexual behavior and nearly 25.6 million Americans (11%) acknowledge at least some same-sex sexual attraction.[xviii] This is not inconsequential

because there is no other evidence in population studies that suggests the impact of this demographic.

The concept of transgender identification is complex, varied, and much of it is beyond the scope of this paper. A demographic profile released within the last month notes that adults who identify as transgender are more racially and ethnically diverse than the U.S. general population.[xix] Extensions to the rest of the world suggest that the US census of LGBT+ may be representative as well. We see as recently as this month how transgender is increasingly breaking barriers with international competitions for beauty contests even in highly conservative and Muslim nations such as Indonesia by crowning Miss Waria (Transgender) Indonesia.[xx]

The US Census: Accuracy and Accountability

The US Census captures data for the LGBT population through the same-sex population households category in the American Community Survey (ACS) and the American Housing Survey (AHS)[xxi], a publication "sponsored by the Department of Housing and Urban Development (HUD) and conducted by the U.S. Census Bureau."[xxii] However, the same-sex households category applies to a wide range of living situations that includes heterosexual and LGBT populations, including but not limited to, family members of the same-sex living together, roommates, co-habiting life partners, and more recently married same-sex couples. This push towards collecting data on the LGBT population was started by the 2013 Supreme Court decision that made Section 3 of the Defense of Marriage Act unconstitutional.[xxiii] This piece of legislation attempted to deny same-sex married couples from receiving federal benefits. Since the act was declared unconstitutional, the Office of Personal Management required data on same-sex married couples to be collected to estimate the beneficiaries for the Federal government.[xxiv]

Despite best intentions to collect accurate information on the same-sex married couples households category, the Census noticed that significant errors existed due to discrepancies in data between the 2013 American Housing Survey and the Social Security Administration (SSA) data based on relationship and gender of the people being asked. Same-sex married couples who identified as such in the 2010 US Census, marked themselves as opposite-sex married couples in the SSA data. However, the errors were less than 1% for opposite-sex married couples who identified themselves as same-sex married couples. This issue is not unique to the US and also occurred in the Census of Canada and also the French census (Enquête annuelle de recensement, EAR).[xxv] The main reason for this mistake is the way the questions were phrased in the AHS.

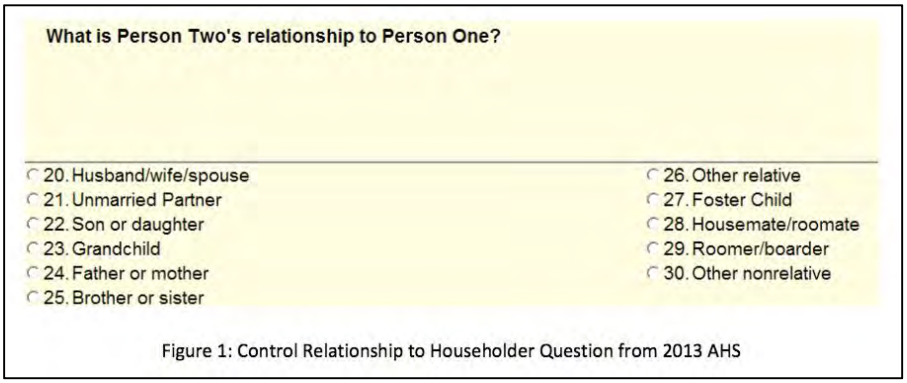


Figure 1: Control Relationship to Householder Question from 2013 AHS

Source: Lewis, J., Bates, N., & Streeter, M. (2015)

The Census Bureau used an indirect approach to measure couple type, including same-sex couples, in the 2010 Census and other most recent Census surveys. Information from both the relationship to householder and sex items is needed to classify couples. Figure 1 shows the control relationship to householder item as it appears in the 2013 AHS. As an example of how this information is used to classify couples, if Person Two's relationship to Person One was recorded as 'Husband/wife/spouse,' and both Person One and Person Two's sex recorded as 'Male,' they would be considered a same-sex married couple.[xxvi]

Due to the confusing nature of the above example, the US Census proposed a new question below that reduces ambiguity between the relationship between "Person One" and "Person Two's" relationship between each other as we can see below in Figure 2.

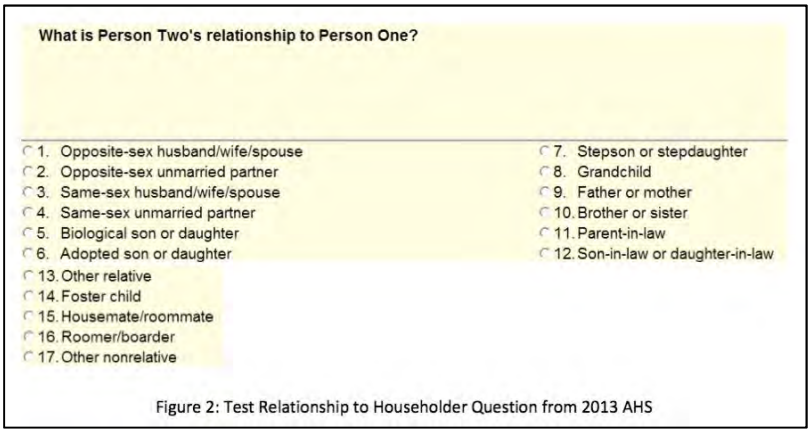
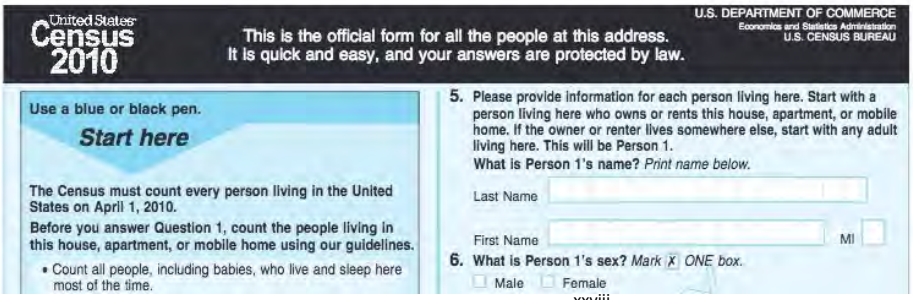


Figure 2: Test Relationship to Householder Question from 2013 AHS

Transgendered individuals represent another potential difficulty in US Census accurate measurement of the LGBT+ communities. The Census recognizes that there are likely transgendered individuals that exist according to a fairly recent analysis of 2010 US Census records against Social Security (SSA) name change records based on the gender of the first name.[xxvii] However, based on the current US Census forms for the last decennial Census and the most recent American Community Survey, there are only "Male" and "Female" offered as choices.



Source: US 2010 Census[xxviii]

Source: 2015 American Community Survey[xxix]

Since the issue of gender and sexuality is complicated it will be interesting to see the direction that the US Census takes to measure the transgender community, should it decide to do so.

Telling the Story: The power of the personal narrative

It appears that nearly every family has a member who is part of the LGBT+ community. The publishing and creative fields have depicted increasingly positive exposures resisting the stereotypes that once were the mainstay of the community. As families have become more open and welcoming of children who are expressing their sexual orientations at earlier ages, the need for public schools to accept all students regardless of circumstance on their terms, the communities have become safer. This would not have happened over the last few decades if support organizations such as PFLAG (formerly Parents and Friends of Lesbians and Gays) had not become so instrumental in educating the public, and fighting acts of intolerance and bullying.

One of the largest milestones of the LGBTQ+ community is the legalization of same sex marriage by the US Supreme Court in June 2015.[xxx] However, even with the legalization of same sex marriage, there has been pushback from individual states that are opposed to same sex marriage for largely religious and fundamentalist reasons.[xxxi]

Despite the victory of same sex marriage legalization, there are still other ways where LGBTQ+ individuals are discriminated. The ACLU (American Civil Liberties Union) is an organization that defends the individual rights and liberties guaranteed to all people in the US.[xxxii] Some of the major issues that they have identified as current issues are the following:[xxxiii]

- LGBT Parenting
- LGBT Relationships
- LGBT Youth
- LGBT Nondiscrimination Protections
- Transgender Rights
- Bullying

Another critical advocacy organization, Lambda Legal, founded in 1973, is the oldest and largest national legal organization whose mission is to achieve full recognition of the civil rights of **lesbians, gay men, bisexuals, transgender people** and those with **HIV** through impact litigation, education and public policy work. Examples of recent work by Lambda Legal have included:

- Influencing prisons to defend and protect incarcerated transgender inmates
- Challenging how fervent religious beliefs have allowed small business owners, from barbers to bakers, etc, to choose their clients and whether to serve them
- Defending guidelines issued by the Obama Administration calling for public schools to respect transgender students by using the right pronouns, combatting bullying and allowing students to use restrooms in line with their gender identity and creating Trans-Student Rights Watch
- Issuing a set of trans-friendly model policies for hospitals to follow, *Creating Equal Access to Quality Health Care for Transgender Patients (http://www.lambdalegal.org/publications/fs_transgender-affirming-hospital-policies)*

Families are sharing experiences that document how children as young as six year olds are confronting their sexuality and communicating that they identify with a gender other than what they are. Requests for cross-dressing and identifying sufficiently to be comfortable to ask for accommodations in school has forced parents to face these issues with younger children than earlier documented. How to offer parental support and love while knowing that their children may suffer as a result of their choices is a parenting skill many need to face. Some families have decided to be "open and honest about it, celebrating these stories instead of hiding it."[xxxiv]

Other families have found how athletics served as the focus of communicating their dilemmas. Sports teams are gender based and when a high school in Santa Monica, California honored its mission of diversity and acceptance and chose to accept a student's wish to play baseball on the boys' team. Originally enrolled as Emma, he became Jake and joined the baseball team. Never being able to find or create the right sisterhood as a girl, experiencing depression and being suicidal, but finally finding kindness at the foundation of his school community, his mother sums it up best as saying, "I would recommend to parents to look in their child's eyes, look at their smile, look beyond any gender and know this is your offspring, this is going to be your child's story, and you're going to want to be part of it."[xxxv]

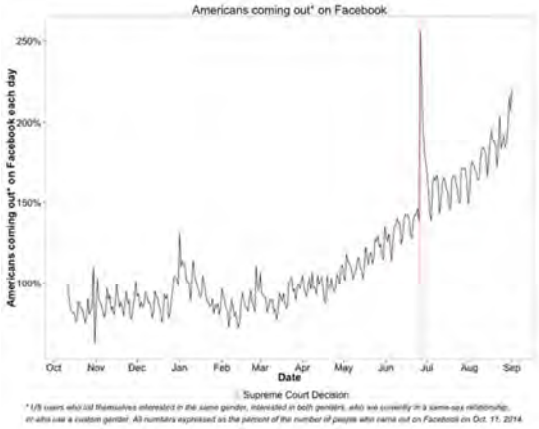
Personal narratives and oral histories became increasingly visible in journals, magazines, plays, music lyrics, photography, film and other distillations of life and were captured to share experiences and reduce the stereotyping and enhance the comfort levels psychologically and emotionally of the gay population. It also bridged the gap of ordinary members of society being able to tell their story without ridicule and humiliation. Obituaries and society pages of local newspapers with growing frequency report on a subject's gender identity, contributing to another level of accuracy, objectivity in contributing to genealogy

and history. Social media on different platforms enables resources such as the "Girls Like Us at the gluproject" to evolve as a mixed media project where Saira Awan explores the stories behind the faces of girls/women/ladies of all ages, races and backgrounds.[xxxvi]

Society's tolerance levels have elevated to where there is critical mass in the general population and a significant proud and visible element for whom total social acceptance is expected and can be shared. However, concerns are raised from the Gay, Lesbian and Straight Education Network's (GLSEN) 2016 report, "Teasing to Torment: School Climate Revisited"[xxxvii] about the immediate need for high school teachers and counselors to become better prepared to respond to and support LGBT+ youth in their schools and to combat the bullying and harassment that students experience and perpetrators are uneducated about.

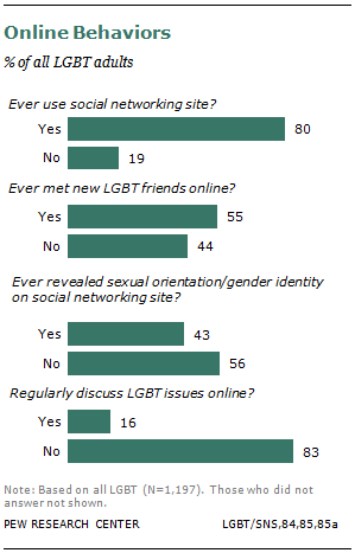
More Social Media and LGBTs

One of the major events that sparked an increase in social media was the events leading up to and also the after effects to the passage of marriage equality in the US. Facebook, recently posted a chart that indicated the number of users "coming out" on their platform as shown in the graphic below. Although the time period only covers a one-year span, we can see a sharp increase through Facebook. We can extrapolate that similar effects happened on other social media platforms as well during this time period.



Source: Contently - How Facebook Changed LGBT Culture, in 3 Charts

The Pew Research Center published a report in 2013 that surveyed LGBT+ Americans, in particular the use of social media sites in their daily life. The results showed interesting results as most of the LGBT adults they surveyed use social networking sites, however the results conflict somewhat with the noted Facebook example in terms of revealing their sexual orientation as the Pew report stated that a slight majority chose not to reveal their orientation.[xxxviii] Part of this may be due to the timing of the report as it occurred two years before the passage of marriage equality in the US. There are scores of online social dating sites and apps for people to connect and meet likeminded friends, many of which offer choice for both LGBT+ and heterosexual members.



Source: Pew Social Trends – A Survey of LGBT Americans

Social media as a whole resonates with the youth of any sexual orientation and the LGBT+ youth is no exception. The "It gets better project" (http://www.itgetsbetter.org/) is dedicated to helping LGBT youth who faced bullying, harassment, and coming out issues that life "gets better" and that the suffering that they currently face is just a small part in their struggle for identity and meaning in their own lives. This resource with social media links and a Twitter feed of @ItGetsBetter shows a glimmer of hope and comfort to those who have little to no support in communities with less visible LGBT presence or youth who want to identify with role models but have none nearby. By providing an avenue for outreach and more social services, helplines, networking and community building, it is probable that this site and others, will ease the suffering of troubled LGBT youth and prevent suicides and the downward spiral of negative social behavior.

Recent history

On Saturday, June 28, 1969, New York City police raided the Stonewall Inn, a bar catering to gay clientele. During this time period, harassment tactics used by police for the LGBT establishments was common. However, on this particular night, the bar patrons and residents surrounding the Stonewall Inn had enough of being harassed and refused to leave the area when directed. This confrontation between protesters and police lead to six days of demonstrations and riots over "homophile rights" in the US. Despite the size and controversial political nature of the event, it was not highly publicized in the press until many years later when films and documentaries depicted its significance. This incident was credited as the "spark" that ignited the gay rights movement in the US. In June 2016, just this year, the US National Park Service commemorated the Stonewall Inn in Lower Manhattan as a National Monument in recognition of the struggle for civil rights for the LGBT community.

Another political milestone for the LGBT community was the AIDS crisis. In the early 1980s, a disease called "Gay Related Immune Deficiency" hit the gay male community in Los Angeles with full force. Gay men started to die of a mysterious illness that scientists were unable to diagnose and treat. The disease was renamed the Acquired Immune Deficiency Syndrome and the emphasis had shifted from being a "gay disease" to a disease that affected all people that were susceptible to coming in contact with the disease through high risk behaviors. Towards the late 1980s and early 1990s, the US government started a public health campaign that controlled the spread of AIDS in the US. Today, the emphasis is on addressing worldwide measures to suppress AIDS and treat the related social and family issues.

Additional Unique Health Concerns

Tobacco use, and illegal drug and substance abuse among the LGBT+ census has proven to exhibit how it is at an increased risk for the adverse effects of tobacco use, given the high prevalence of use, especially smoking among this population. This evidence-based data

suggests how LGBT serving organizations should implement more aggressive tobacco dependence and substance abuse treatments that will encourage diversions from such damaging health conditions.[xxxix] Other health issues in addition to AIDS have challenged this community. Sex-change surgeries do not come without complications and are carried out via numerous procedures. Mental health services for this population vary widely and access is often dependent on individual economic abilities and local social services that vary greatly.

The Gay Press & Publishing: Capturing for the future

Two major newspapers have provided coverage and advocacy for the LGBT+ community and are based in Los Angeles and released by Regent Communications. *The Advocate*, a bi-monthly magazine issued from 1967 captures news, stories, and interviews and is a platform for social and political advocacy with a subscription of just shy of 200,000. *Out,* more of a lifestyle and consumer magazine for LGBT+ readers began publishing in 1992 and has a current subscriber base of more than 200,000. These outlets provide channels for marketing and advertising to this growing readership and along with 8 other trade publications all gaining traction.[xl]

The academic and scholarly resources that track and explore LGBT+ society have proliferated in recent years. The One Archive located at the University of Southern California Libraries claims to be the largest repository of LGBTQ materials in the world with beginnings in 1952.[xli] Subject access to resources and collections has been improved by a February 2016 update of the Library of Congress Subject Headings.[xlii] Most major publishers, the commercial trade and scholarly presses have released many series of books and journals that are both interdisciplinary and couched in specific subject fields. The citation landscape has imploded with a growth of traditional resources and a recent study published in 2012 demonstrates how from 1974-2010 the materials remain more mainstream and academic rather than popular or community and activist sources.[xliii] Databases such as Gender Studies Database, GenderWatch, Defining Gender, LBGT Life with Fulltext[xliv] and other databases that cover history, politics, women's studies, psychology and sociology all have relevant content sourced, cited and indexed from a variety of primary and secondary materials.

One of the most important resources to launch this year was the Digital Transgender Archive,[xlv] a culmination of contributions from more than 20 nonprofit agencies, libraries and institutions which are collaborating to create the world's largest transgender history digital archive. The College of the Holy Cross, a Jesuit institution is leading the project with Professor K.J Rawson as director and archivist. Academe has a first endowed chair of transgender studies in Professor Aaron Devor at the University of Victoria created by the generosity of billionaire Jennifer Pritzker of Chicago. Devor and Rawson hope the Digital Transgender Archive "will provide a more nuanced look at trans-history, and believes the research has the power to enact policy changes that help transgender people."[xlvi] One of the largest collections of grey literature this archives possesses more than a football field of archival boxes full of memorabilia, papers, correspondence, audio tapes, photos, programs, full runs of newsletters, earlier tabloids and journals that document the history of transgender experiences globally, including the forty plus year history of the Fantasia Fair, the longest running conference for transgender people held in Provincetown since 1975. The Trans Lives Matter movement has also become more visible due to the added exposure that the Black Lives Matter has generated over the last two years and Black Transgender individuals have found community and support in this community.

Grey literature has been a hallmark for different threads of gender studies reflecting mens' studies, womens' studies and trans' studies. Global coverage has also proliferated from very open cultures to those that have been less than hospitable to LGBT+ populations. Religion and politics contribute to that lack of transparency but are not the only causes. Within more of an arts and literary focus, *Vetch* captures trans-poetry and poetics since its inaugural release in 2015. Independent bookstores in urban centers and college towns were known to concentrate on offering a range of literature that served these interests. Online commerce has an example of a digital storefront (https://www.lgbtbookshop.com/) specializing in print and eBooks for and about the LGBT+ lifestyle.

Additional resources that showcase the history and contemporary lives of LGBT+ include fanfiction produced online for gay communities, participative blogs to which readers can

contribute, comment and gain insights from, photo archives captured by museums and libraries and books such as the recently released *On Christopher Street: Transgender Stories*.[xlvii] There are many social media installations via Facebook, Twitter, Instagram, Tumblr, and others that capture lifestyle stories such as those revealed and shared via #GirlsLikeUs[xlviii]. Today there are many museums worldwide, including in the US, the LGBT History Museum in San Francisco in the heart of the Castro, in Washington, DC, the Smithsonian and its newest facility, the National Museum of African American History and Culture[xlix] that have rotating exhibits that chronicle the LGBT+ community's rich history, but today there are 9-10 national monuments and museums dedicated and focused on preserving the past and collecting artifacts that document sexual minorities.[l]

Sexuality Archives & Psychiatric Findings

Beginning in the 1940s, Alfred Kinsey documented the lesbian and gay subculture in *Sexual Behavior in the Human Male.* His research demonstrated that same-sex behavior was not all the same by disputing that all adults were either homosexual or heterosexual and that there was a population best described as "sexually fluid." The American Psychiatric Association removed homosexuality from its official *Diagnostic and Statistical Manual of Mental Disorders* (DSM-II) in 1973. This decision occurred in the context of momentous cultural changes brought on by the social protest movements of the 1950s to the 1970s beginning with the Civil Rights movement, then evolving on to the women's and gay rights movements. Having successfully challenged the police and government attempts to shut down public places where gay people gathered, gay activists would soon challenge psychiatric authority as well.

Before the Stonewall riots, homophile groups had accepted the medical view of homosexuality as a mental disorder. Their view had been that accepting homosexuality as disease meant treating it as a disability, rather than a moral or religious sin, and would lead to more objective and humane attitudes. The DSM-II diagnosis of Sexual Orientation Disturbance (SOD) replaced Homosexuality. Accordingly, individuals comfortable with their homosexuality were no longer considered mentally ill. Only those who were "in conflict with" their sexual orientation had a mental disorder (SOD). This compromise engendered continued controversy. Those opposing it pointed out there were no reported cases of unhappy heterosexual individuals seeking treatment to become homosexual. This problem was addressed in the revisions of the 1980's DSM-III where SOD was replaced by ego-dystonic homosexuality (EDH) and did not come without enormous controversy. The DSM-IV does not include homosexuality per se as a disorder, but still permits the diagnosis of "Sexual Disorder Not Otherwise Specified" for someone with "...persistent and marked distress about sexual orientation," while one can conclude that this latest edition was more reliant on empirical data and less influenced by the politics of the day.

Public Awareness

The 2016 US presidential election was an important event for LGBT+ politics as Democratic candidate Hillary Clinton supported upholding marriage equality for LGBT+ Americans as was stated on her campaign website[li] and Donald Trump took the opposing view by pledging to repeal marriage equality[lii] but has since accepted that it has become part of the domestic landscape and will not change it. Internationally same-sex marriage laws have expanded throughout much of Europe, however there remains staunch resistance in Asia, until the current news that in 2017 Taiwan may have sufficient votes in its legislature to pass such legislation becoming a pioneer on that continent.[liii] President-Elect Trump, also mentioned that he will "defend" LGBT+ Americans as a reaction from the Orlando nightclub massacre.[liv] Even the Log Cabin Republicans, a LGBT+ group allied with the Republican Party, did not endorse Donald Trump for president because of his anti-LGBT stance on marriage equality.[lv] With the victory of marriage equality in the US, the battleground for LGBT+ political issues shifted from the national spotlight to the individual states for LGBT+ issues. The Human Rights Campaign, a LGBT+ civil rights organization that advocates for LGBT+ rights, maintains a map of State Laws and Policies that outlines the different laws from each state.

Source: Human Rights Campaign

This shift towards state politics rather than national issues makes the individual state leadership candidates more important based on their views of LGBT+ issues. Also to complicate matters, as mentioned previously, LGBT+ individuals not only identify with their own sexuality but also maintain intersectionalities with their race/ethnicity, socioeconomic status, and other areas in their life that may influence the political viewpoints of the voter.

Conclusion: Leveraging diversity

Since we began work on this paper, much has happened in the United States with the outcome of the national election being one we could not have predicted. The media has spun many scenarios about how the results happened. One political scientist, Joseph Nye helps us understand that there are two kinds of presidencies. Those that are transformational and those that are transactional. The former seeks to change the political landscape in fundamental ways while the latter seeks to manage the landscape much as it is.[lvi] We have yet to see where on the spectrum Mr. Trump will fall. The Equality Act when passed will amend the Civil Rights Act of 1964 to include protections that ban discrimination on the basis of sexual orientation, gender identity and sex in the areas of employment, housing, public accommodations, public education, federal funding, credit, and the jury system, and is critical since LGBT+ Americans in 30 states live without fully inclusive nondiscrimination laws. We do know that since the election there has been a "big uptick in incidents of vandalism, threats, intimidation," and expressions of hate, intolerance and fear that major human and civil rights may be removed. We have a Vice-President-Elect who last year as governor of Indiana, signed the "Religious Freedom Restoration Act," which allows companies to discriminate against LGBT+ employees and customers based on firmly held religious beliefs. And Mr. Pence proposed cutting funding for HIV treatment and using the money for "gay cure therapy," killing off Planned Parenthood, and together with Mr. Trump will nominate Supreme Court Justices who will push the Christian Fundamentalist persuasion and potentially reverse Roe v. Wade and the ability for women to elect to have abortions. Bodies of Pence's association such as the American Conservative Union, the appointment of a Chief Strategist leading the the pull toward the "alt-right" where white supremacists and nationalists are getting air time are disturbing and driving mainstream members of society to protest and worry about their future. Populations already marginalized within the LBGT+ community had every reason to be concerned as it remains unclear whether the Republican platform would reinstate "don't ask, don't tell" regarding service in the military, support for state laws limiting which public bathroom transgender people can use and support for a parent's right to subject gay and transgender children to conversion therapy to change their sexual orientation or gender identity.[lvii] It will be an uneasy time these next few months, but we remain optimistic that the implications of going backward would set a very bad precedent and create a dark cloud under this administration before they even have a chance to try their hand at fixing healthcare, redirecting the economy and creating jobs and opportunities for all. The National Center for Transgender Equality has championed the causes important to their community and has just released its findings from the second U.S. Trans Survey conducted last year that garnered 27,000 responses. Its predecessor released in 2011 with 660 respondents is the most cited study

about transgender people. This current survey demonstrated regarding the bathroom controversy that:

> **59% have avoided bathrooms** in the last year because they feared confrontations in public restrooms at work, at school, or in other places.
>
> **12% report that they have been harassed, attacked, or sexually assaulted** in a bathroom in the last year.
>
> **31% have avoided drinking or eating** so that they did not need to use the restroom in the last year.
>
> **24% report that someone told them they were using the wrong restroom** or questioned their presence in the restroom in the last year.
>
> **9% report being denied access to the appropriate restroom** in the last year.
>
> **8% report having a kidney or urinary tract infection**, or another kidney-related medical issue, from avoiding restrooms in the last year.[lviii]

A growing concern is the increase in reporting about the elderly regardless of what population of LGBT+ they represent. For those who have been distanced from family and who have experienced poverty and life under the radar, growing old and destitute has been reaffirmed as they find themselves often in unwelcome public residential communities where bias, signs of harassment and abuse have driven residents to file lawsuits against retirement facilities. If successful the current Wetzel lawsuit may establish legal precedent for the responsibility of "housing providers to actively address discrimination based on gender identity and sexual orientation under the federal Fair Housing Act where the law states that discrimination based on "sex" is prohibited."[lix]

There is much uncertainty, much angst and after accomplishing so much for human rights, leveraging grey is no longer a sufficient hue but clarity about equality for all citizens is essential. It is about legitimizing life whether it be any part of the LGBT+ population.

References

[i] Stone, A, (2015) Gay and Lesbian Movements. *International Encyclopedia of the Social & Behavioral Sciences*, 2d ed, vol. 9: 632.

[ii] Wickman, F. (2013). A Rainbow Marriage: How did the rainbow become a symbol of gay pride? *Slate Magazine*, June 26. http://www.slate.com/articles/life/explainer/2012/06/rainbows_and_gay_pride_how_the_rainbow_became_a_symbol_of_the_glbt_movement_.html

[iii] Greenblatt, A. (2015, December 11). Transgender rights. *CQ Researcher*, *25*, 1033-1056. Retrieved from http://library.cqpress.com/

[iv] Intersectionality - definition of intersectionality in English. Oxford Dictionaries. (n.d.). Retrieved October 03, 2016, from https://en.oxforddictionaries.com/definition/intersectionality

[v] Wikipedia, *https://en.wikipedia.org/wiki/**LGBT***

[vi] Lillian Faderman (2015). *The Gay Revolution: The Story of the Struggle*. NY: Simon & Schuster, xix-xx.

[vii] Human Rights Campaign. (n.d.). Glossary of terms. Retrieved November 06, 2016, from http://www.hrc.org/resources/glossary-of-terms

[viii] Kornhaber, S. (2015). Modern Family Effect: Popular Culture's Role in the Gay Marriage Revolution. Atlantic Monthly, June 25. http://www.theatlantic.com/entertainment/archive/2015/06/gay-marriage-legalized-modern-family-pop-culture/397013/

[ix] Woods, G. (2016). *Homintern: How Gay Culture Liberated the Modern World*. New Haven, CT: Yale University Press, xi.

[x] *Ibid:* xi-xii.

[xi] GLADD (2016). 2016 *Where we are on TV*. Retrieved from http://www.glaad.org/wherewereontv15

[xii] GLADD (2005). 2005 *Where we are on TV*. Retrieved from http://www.glaad.org/publications/tvreport05

[xiii] International Transgender Day of Remembrance, https://tdor.info/about-2/

[xiv] Ziri, D. (2016). Being transgender and Jewish. *The Jerusalem Post*, August 6.

[xv] The Evangelical Network, http://www.theevangelicalnetwork.net/

[xvi] Muslims for Progressive Values, http://www.mpvusa.org/mpv-principles/

[xvii] Gary J. Gates (2011). How many people are lesbian, gay, bisexual and transgender? Williams Center Report. http://williamsinstitute.law.ucla.edu/wp-content/uploads/Gates-How-Many-People-LGBT-Apr-2011.pdf

[xviii] Ibid.

[xix] Andrew R. Fores, Taylor N.T. Brown & Jody L. Herman (2016). Race and Ethnicity of Adults who Identify as Transgender in the United States. http://williamsinstitute.law.ucla.edu/wp-content/uploads/Race-and-Ethnicity-of-Transgender-Identified-Adults-in-the-US.pdf

[xx] Trans Pageant Held in Secret (2016). *Los Angeles Times*, November 15: A4.

[xxi] Lofquist, D. (2011, September). *Same-sex couple households, American Community Survey Briefs* (United States, US Department of Commerce, US Census). Retrieved October 9, 2016, from http://www.census.gov/prod/2011pubs/acsbr10-03.pdf
[xxii] United States, US Department of Commerce, US Census. (n.d.). *Data on same sex couples*. Retrieved October 9, 2016, from http://www.census.gov/hhes/samesex/data/
[xxiii] Defense of Marriage Act, H.R. 3396, 104th Cong., (1996).
[xxiv] Lewis, J., Bates, N., & Streeter, M. (2015). *Measuring same-sex couples: the what and who of misreporting on relationship and sex* (SEHSD Working Paper 2015-12). Washington DC: US Census.
[xxv] Ibid.
[xxvi] Ibid.
[xxvii] Harris, B. (2015). *Likely Transgender Individuals in U.S. Federal Administrative Records and the 2010 Census* (CARRA Working Paper #2015-03). Washington DC: US Census.
[xxviii] United States, US Department of Commerce, US Census. (2008, September 25). *2010 United States Census Form*. Retrieved October 9, 2016, from https://www.census.gov/schools/pdf/2010form_info.pdf
[xxix] United States, US Department of Commerce, US Census. (2014, June 17). *2015 American Community Survey*. Retrieved October 9, 2016, from http://www2.census.gov/programs-surveys/acs/methodology/questionnaires/2015/quest15.pdf
[xxx] Liptak, A. (2015, June 26). Supreme Court ruling makes same-sex marriage a right nationwide. *New York Times*. Retrieved October 3, 2016, from http://www.nytimes.com/2015/06/27/us/supreme-court-same-sex-marriage.html?_r=0
[xxxi] Wetzstein, C. (2015, July 6). States rebel against Supreme Court gay marriage ruling. *The Washington Times*. Retrieved October 3, 2016, from States rebel against Supreme Court gay marriage ruling
[xxxii] ACLU FAQs. (n.d.). Retrieved October 03, 2016, from https://www.aclu.org/faqs#1_1
[xxxiii] ACLU LBGT Rights. (n.d.). Retrieved October 03, 2016, from https://www.aclu.org/issues/lgbt-rights
[xxxiv] Frankel, L. (2016). From He to She in First Grade. *The New York Times*, September 18: ST:6.
[xxxv] Plaschke, B. (2016). Transcending the game: Playing baseball for his Santa Monica prep school helps transgender teen convey message of hope and acceptance. *Los Angeles Times*, April10: B1.
[xxxvi] https://twitter.com/gluproject
[xxxvii] https://www.glsen.org/sites/default/files/TeasingtoTorment%202015%20FINAL%20PDF[1]_0.pdf
[xxxviii] Pew Research Center. (2013, June 13). A Survey of LGBT Americans. Retrieved November 07, 2016, from http://www.pewsocialtrends.org/2013/06/13/a-survey-of-lgbt-americans/
[xxxix] Less, J.G.L, et all, (2014). Promotion of Tobacco use Cessation for Lesbian, Gay, Bisexual and Transgender People: A Systematic Review. *American Journal of Preventive Medicine* 47 (6):823-881.
[xl] www.cision.com/us/2012/06/top-10-gay-and-lesbian-magazines/
[xli] http://www.onearchives.org/
[xlii] "Library of Congress Subject Headings for Lesbian, Gay, Bisexual, Transgender, Queer (LGBTQ) Topics" compiled by John De Santis (2000), Thomas Bolze and Deborah LaFond (2010), last revised February 2016. http://library.albany.edu/subject/guide/lgbtqi
[xliii] Antell, K.(2012). The Citation Landscape of Scholarly Literature in LGBT Studies: A Snapshot for Subject Librarians. *College and Research Libraries*, November, 29: 584-602.
[xliv] https://www.ebscohost.com/academic/gender-studies-database, http://www.proquest.com/products-services/genderwatch.html, http://www.gender.amdigital.co.uk/, https://www.ebscohost.com/academic/lgbt-life-with-full-text
[xlv] https://www.digitaltransgenderarchive.net/about/overview
[xlvi] Garfield, L (2016). Inside the World's Largest Transgender Archive. *Business Insider*, March 17. www.businessinsider.com/inside-the-worlds-largest-transgender-archive-2016-3/
[xlvii] Seliger, M. and Mock, J. (2016). *On Christopher Street: Transgender Stories*. New York: Rizzoli.
[xlviii] http://janetmock.com/2012/05/28/twitter-girlslikeus-campaign-for-trans-women/
[xlix] Dupuis, A. (2016). New Smithsonian Places African American LGBT Community In Full View. *Insight Magazine*, August 23. http://instinctmagazine.com/post/new-smithsonian-places-african-american-lgbt-community-full-view
[l] https://www.bustle.com/articles/175638-9-lgbt-museums-memorials-around-the-world-you-should-visit
[li] Hillary Clinton. (n.d.). Hillary Clinton on LGBT rights and equality. Retrieved November 07, 2016, from https://www.hillaryclinton.com/issues/lgbt-equality/
[lii] Fox News Network. (2016, January 31). Ted Cruz attacks Donald Trump's financial record; Trump responds. Retrieved November 07, 2016, from http://www.foxnews.com/transcript/2016/01/31/ted-cruz-attacks-donald-trump-financial-record-trump-responds/
[liii] Taiwan as a Same-Sex Marriage Pioneer. Editorial, *New York Times*, November 26, 2016: A20.
[liv] Donald Trump. (n.d.). Why LGBT Community Should Back Trump. Retrieved November 07, 2016, from http://www.donaldjtrump.com/media/why-lgbt-community-should-back-trump
[lv] NPR. (2016, October 30). 2016 Has Been A Mixed Bag For LGBT Politics. Retrieved November 07, 2016, from http://www.npr.org/2016/10/30/499895397/2016-has-been-a-mixed-bag-for-lgbt-politics
[lvi] McManus, D. (2016). Which Trump will show up in D.C.? *Los Angeles Times*, November 13: A20.
[lvii] Stack,L, (2016). Trump Win Seen as 'Devastating Loss' for Gay and Transgender People. *The New York Times*, November 11: P9.
[lviii] U.S. Transgender Survey, 2016 Preliminary Findings (http://www.ustranssurvey.org/preliminary-findings)
[lix] Miller, M. (2016). Fighting Gay Bias at a Rest Home. *The New York Times*, November 19: B1,4.

Managing Diversity in the International Nuclear Information System

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

Although diversity is defined as the state of having many different forms, types, ideas or properties, most often it is associated with cultural diversity or different ethnic backgrounds. Biology, religion, and political science researchers have their own view of diversity, such as biodiversity (i.e. variety of life on earth); religious pluralism (i.e. multi-confessionalism, multi-faith), or various shades of democracy (i.e. freedom of opinion or expression). Diversity is also applicable to information management, usually through the diversity of information resource formats and the variety of information users (i.e. customers or clientèle).
This paper adopts a holistic approach to information management, in particular to the management of scientific and technical information (STI), as carried out by an international repository, and attempts to identify various intrinsic and extrinsic properties which include elements of apparent diversity. In addition to the above mentioned diverse information formats and the variety of information users, intended and actual, this paper also examines the diversity of subject content, temporal distribution, geographic coverage, variety of distribution channels, search paths, and composition of the staff directly involved with the running and management of an information repository.
The data used in this research comes from the International Nuclear Information System (INIS), which has collected, processed and provided access to more than 4 million bibliographic references of publications, documents, technical reports, non-copyrighted materials, and other grey literature, as well as over a million full-texts. The INIS repository represents one of the world's largest collections of published information on the peaceful uses of nuclear science and technology. The history of INIS throughout the last 47 years represents a wealth of information on the successful management of diversity, starting with 154 member states and international organization who share and allow access to their valuable nuclear information resources while preserving them for future generations, offering nuclear information repository used annually by millions of scientists, researchers, engineers, technicians, students, managers, and government employees.

Keywords: nuclear information; information management; document repository; grey literature; diversity; INIS; IAEA

Introduction

The main topic of this paper is leveraging diversity in grey literature. However, as we can see from the title, there are two major concepts here put into one topic. The first one is ***diversity*** and the second one is ***grey literature***. Taking this into consideration, the first part of the paper concentrates on the review of conventional diversity, while the second part focuses on information diversity, more specifically on diversity related to grey literature and the way it is being managed in the International Nuclear Information System (INIS)[1].
Conventional diversity will be presented as a concept, followed by a review of challenges, and culminating with a look at values and benefits. The review of diversity in INIS will start with identifying various aspects of diversity, followed by diversity-based results, and concluding with ways that diversity is being managed.
In closing, some general observations and conclusions will be drawn. It is expected that examples given throughout will demonstrate potential strength and benefits that can be achieved through diversity.

[1] The International Nuclear Information System (INIS) hosts one of the world's largest collections of published information on the peaceful uses of nuclear science and technology. INIS is operated by the International Atomic Energy Agency (IAEA) in collaboration with 154 Member States and international organizations. There are over 4 million bibliographic references to publications, documents, technical reports, non-copyrighted documentation and other 'grey literature', as well as 350,000 full-texts. INIS offers free and open online access to this unique collection of non-conventional literature through its search application (http://inis.iaea.org/search/).

In diversity there is beauty and there is strength - *Maya Angelou*

Grey literature

There are several definitions of grey literature, the most common being the so-called "Luxembourg definition," which was discussed and approved during the Third International Conference on Grey Literature in 1997: "[Grey literature is] that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers." In 2004, at the 6th Grey Literature conference in New York, a postscript was added for purposes of clarification "...not controlled by commercial publishers, i.e., where publishing is not the primary activity of the producing body" (Schöpfel, 2010). As a synonym for grey literature, INIS uses the term 'non-conventional literature' or NCL, which will be mentioned later in this paper.

Diversity concept

Diversity is usually defined as the state of having many different forms, types, ideas or properties[2]. It encompasses acceptance and respect.

Diversity is most often associated with cultural diversity or different ethnic backgrounds. Keeping in mind, diversity also refers to human qualities that are different from our own and those of groups to which we belong; but that are manifested in other individuals and groups. Biology, religion, and political science researchers have their own view of diversity, such as biodiversity (i.e. variety of life on earth); religious pluralism (i.e. multi-confessionalism, multi-faith), or various shades of democracy (i.e. freedom of opinion or expression).

There are many papers written on these types of diversity, but not so many on 'information diversity'. For example, Google Search offers only 42,000 hits for 'information diversity' and over 11 million hits for 'cultural diversity'.

Diversity is also applicable to information management, usually through the diversity of information resource formats, the variety of information users (i.e. customers or clientèle), or a number of different information services.

Diversity challenges

Diversity comes with a number of inherited challenges. The most common ones are stereotypes, bias, generalizations, beliefs and resistance to change.

A ***stereotype*** is a generalized view or popular belief about a particular group of people, usually inaccurate and sometimes offensive (Kanahara, 2006). As the picture on the right shows, we can easily assume the nationality of the character, its characteristics, and other attributes. However, that might be completely wrong and coincidental. A very important lesson here is to connect with the individual, not with the stereotype and not by applying a label, whether positive or negative. Stereotyping is most often related to people or nations, but it can also be related to ideas and things. It should be kept in mind that stereotyping is not necessarily something wrong. It could be quite useful, since it allows us to condense information and make quicker decisions.

Another example of stereotypes is age. This is a big challenge, especially when dealing with and trying to benefit from diversity in the work place. Some research suggests that age related stereotypes bring particularly unfavourable consequences for women and minorities (Shultz & Adams, 2007). A common stereotype related to age is that older generations have poor IT skills, which is often not correct. Being 'too young' or 'too old' is often seen as negative, rather than an opportunity to bridge a gap between generations. Gender is another example of stereotypes, reflected in unequal pay, different career development opportunities, female vs. male jobs, etc.

Age stereotype
- Less motivated
- Not interested in training
- Resistant to change
- Not trusting
- Having health problems
- Vulnerable to work-family conflicts

(Ng & Feldman, 2012)

[2] Merrian-Webster Online Dictionary. https://www.merriam-webster.com/dictionary/diversity

Bias and stereotype are somewhat related. While stereotype is a preconceived idea that attributes certain characteristics to all the members of a class or set, bias is an individual or personal preference[3]. For example, a statement such as - I think that grey literature has no academic or scientific merit - is a stereotype. However, if we have in front of us two equally valid documents, one published in a journal and the other in a personal blog, and we decide to use only the one from the journal – that is bias.

A ***belief*** is an internal feeling that something is true, even though that belief may be unproven or irrational. For example, I believe that tweeting is the best way to quickly exchange information, although that is probably far from the truth. However, changing beliefs is a very difficult task. Often, in order to change beliefs, people need to change their experiences first.

Resistance to change and resistance to diversity are attributed, in part, to the natural fear of change, which often brings with uncertainty, discomfort and loss of control. There are some who outright oppose the concept of diversity (e.g. We've always done it this way), but in practice, this resistance usually manifests itself as a challenge expressed at the very beginning of a diversity based project or program. Some examples of diversity resistance include:

- discrediting information provided by people with diverse backgrounds;
- unwillingness to acknowledge and recognize the contributions of people with diverse backgrounds;
- delaying consideration/implementation of diversity issues;
- rejecting diversity as being too time consuming or complex.

In addition, the frequency of change undergone by a person at the work place is connected to the frequency of the negative emotions reported, which leads to a lowering of trust, lack of engagement, and a decrease in performance (Rafferty & Griffin, 2008). In order to achieve efficient change, all involved need to fulfil the commitments undertaken and not have hidden interests (Grama & Todericiu, 2016).

Diversity values and benefits

It is evident that diversity has very strong ethical, personal, business and social implications.

Ethics – a code of moral principles that set standards of conduct about what is right and wrong, moral and immoral, fair and unfair, and proper and improper in one's behaviour (Seid & Venkataram, 2016). Globalization and an increasingly interconnected world have created tension between universal ethics and local values and norms. What one culture perceives as an ethical action, another culture may not. Organizations need to invest resources and time to train their staff and leaders to properly manage diversity and make ethical decisions. Diversity training is often used as a tool to improve current situations in an organization.

From a ***personal*** aspect, diversity can create curiosity. By being exposed to different people and cultures, we may want to learn more about a particular group which can give us insight into how and why things work as they do and possibly a way to change them for the better.

In ***business***, employing a diverse workforce enables the of use a wider range of talents and skills. This leads to creativity and innovation. Businesses need to mirror the communities and cultures in which they work in order to understand and anticipate the diverse needs of their customers. In ***society***, diversity brings richness and variety. There are always new and interesting things to learn from each other.

Diversity at INIS

When dealing with aspects of diversity within INIS, a number of facets come to mind. Namely, diversity can be related to subject content, information formats, temporal distribution, geographic coverage, distribution channels, and also to the representation of INIS staff. Striving on its diversity, INIS throughout its history has achieved a number of

[3] A list of cognitive biases can be found at http://rationalwiki.org/wiki/List_of_cognitive_biases

remarkable milestones, marking its commitment to the efficient and effective collection, processing and dissemination of nuclear information, as well as its contribution and support for open access to scientific information, with global benefits.

Subject content of the INIS repository covers around 50 well defined categories (IAEA, 2010) which are regularly maintained by INIS, and provides scope descriptions used by national and regional centres to categorize nuclear literature for INIS input. The INIS Joint Reference Series publications are also available on the INIS website.

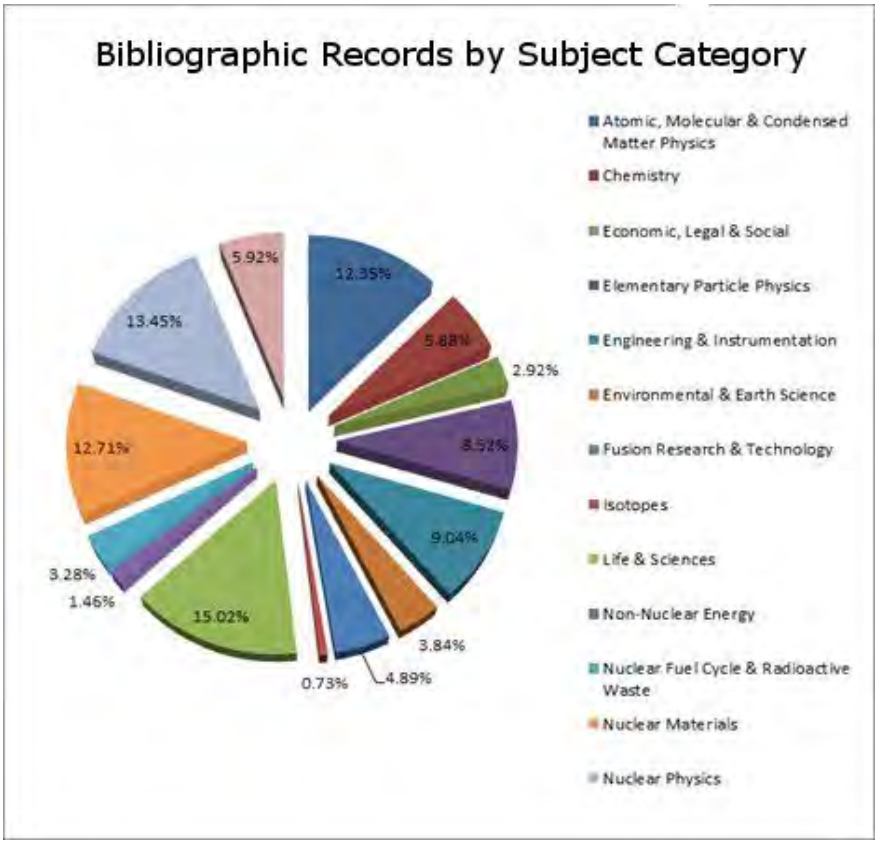


Figure 1: Bibliographic Records by Subject Category

The INIS repository covers all aspects of the peaceful uses of nuclear science and technology, such as nuclear reactors, reactor safety, nuclear fusion, applications of radiation and radioisotopes in medicine, agriculture, industry and pest control, as well as related fields of nuclear chemistry, nuclear physics and materials science. Special emphasis is placed on the environmental, economic and health effects of nuclear energy. Legal and social aspects associated with nuclear energy are also covered. Figure 5 lists a complete set of INIS Subject Categories.

Regarding ***information formats***, the INIS repository consists of seven types of literature – computer media, patents, books, reports, journal articles, miscellaneous and audio-visuals. The main purpose and goal of INIS is to collect non-conventional or grey literature on nuclear science and technology from around the world. This includes various technical reports, project documents, feasibility studies, government reports and any other publications which are not commercially available. However, as shown in the figure below, the INIS repository also contains a number of other forms of conventional literature. Still, it should be noted that while the INIS repository includes links to commercially available journal articles, a great majority of full-texts are, in fact, grey, or non-conventional literature.

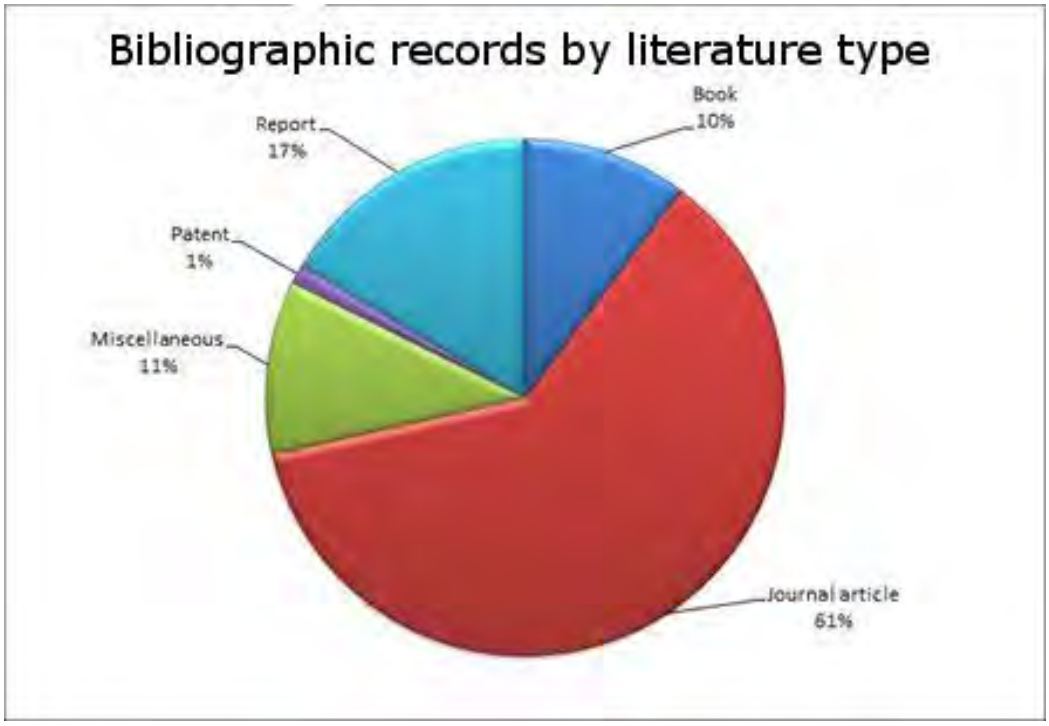


Figure 2: Bibliographic Records by Literature Type

INIS repository

Conventional literature 71%
- Journal articles 61%
- Books 10%

Non-conventional Literature 29%
- Reports 17%
- Miscellaneous 11%
- Patents 1%

Temporal distribution of records in the INIS repository is another aspect of its diversity. As of 31 December 2016, the INIS repository contained 4,000,361 bibliographic metadata records. Of those, 1,093,004 are full-text documents, 761,910 of which are directly available from INIS, with the remaining 331,094 NCLs available from other sources. Only a small portion of the full text documents is restricted and kept for internal use.
During the last 10 years, INIS has added to its repository over 120,000 bibliographic records and 13,000 full text PDF documents annually. The repository is accessible from the INIS website (www.iaea.org/inis).

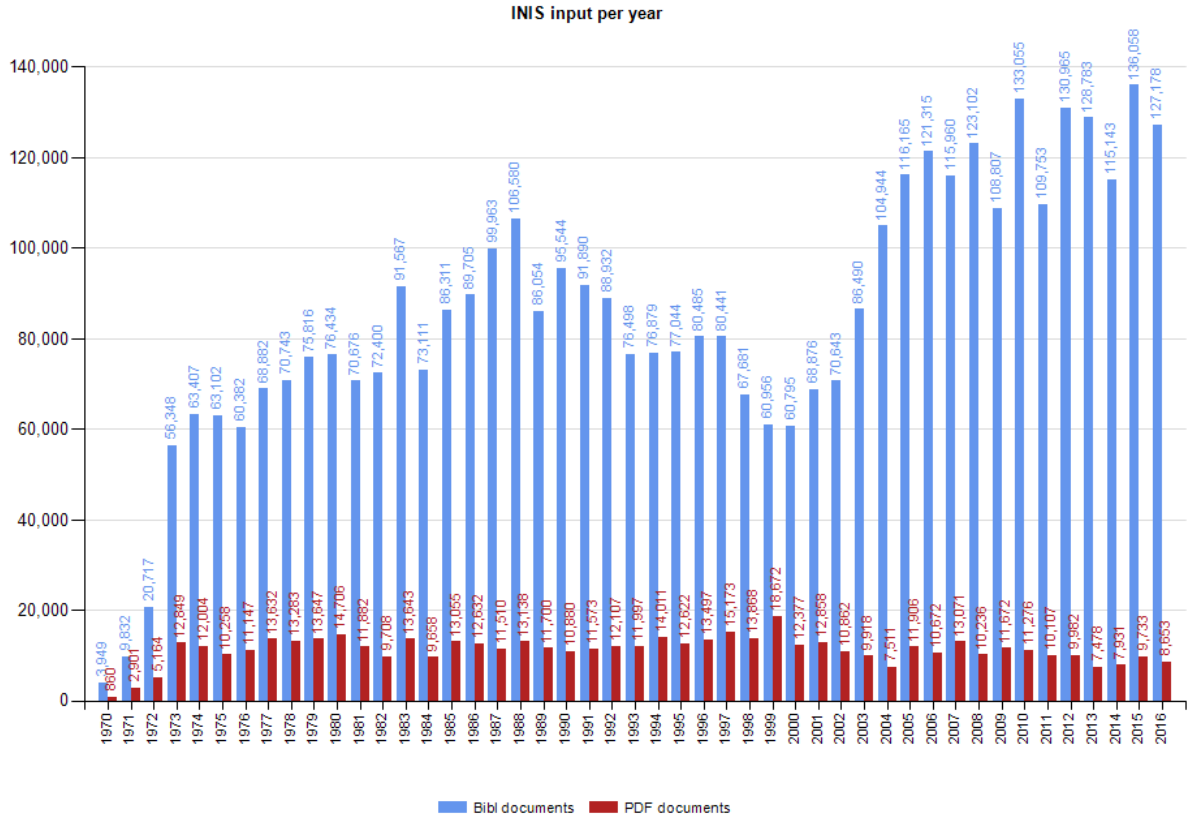


Figure 3: INIS Input per Year

INIS ***geographical coverage*** is almost universal. Currently it includes 150 countries and 24 international organizations. The figure below shows the growth of INIS membership since its creation in 1969.

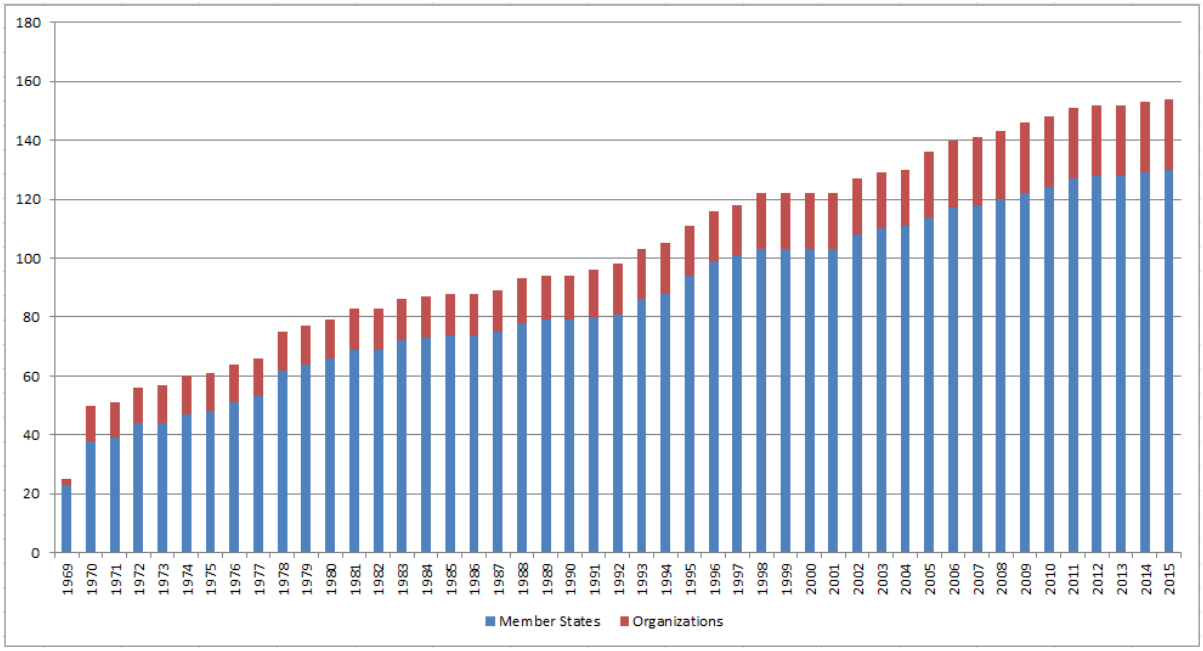


Figure 4: INIS Membership Growth

Besides the geographical coverage of input, that is much diversified as shown above, access and usage of the INIS repository is even more diversified. According to INIS statistics obtained from Google Analytics, users from 229 countries and territories (as defined by

Google) accessed the INIS repository in 2016. A list of the 10 countries with the most visits is given below.

	Country	Sessions	% Sessions
1.	United States	233,928	18.06%
2.	India	137,130	10.59%
3.	United Kingdom	68,701	5.30%
4.	Japan	55,635	4.30%
5.	Germany	51,724	3.99%
6.	South Korea	47,543	3.67%
7.	Canada	43,037	3.32%
8.	France	41,748	3.22%
9.	Brazil	35,167	2.72%
10.	Iran	33,513	2.59%

Figure 5: Visits to INIS

Distribution channels for information and documentation available from INIS include the INIS repository search, which is the main channel, as well as Google and Google Scholar, WorldWideScience.org and various widgets developed for use by outside websites.
IAEA ***staff*** represent a world-class, diverse and multi-disciplinary workforce of more than 2,500 employees working in a wide range of nuclear and related disciplines, coming from over 100 countries. The Nuclear Information Section of the IAEA is made up of 30 staff members from 17 countries and 5 continents.

Since its creation in 1969, INIS progress has been marked by a number of very important ***milestones***, which have left a mark on INIS, on the work of the IAEA, and on the increasing importance of grey literature and nuclear information. Some major milestones are listed below.

1970 INIS - the first nuclear database at the IAEA
1979 INIS - the first IAEA database with online access
1991 INIS & AGRIS - the first IAEA databases on CD-ROM
1996 INIS Web Site - the first Web Site at the IAEA
1998 INIS - the first IAEA database available on the Internet
2009 All Internet users given free and open access to the INIS database
2009 INIS database becomes accessible via WorldWideScience.org
2010 INIS offers the first repository search widget
2011 INIS launches new web search interface using Google-based technology
2012 INIS multilingual Thesaurus integrated with the INIS repository search
2012 INIS repository search includes the IAEA Library catalogue
2013 Browse INIS repository by Subject Category
2014 INIS repository becomes searchable through Google and Google Scholar
2015 Launched Open Nuclear Information eXchange System (ONIXS)
First harvesting program based on Open Access Interface-Protocol for Harvesting Metadata (OAI-PMH)

In 2009, INIS announced that access to the INIS database was open to Internet users around the world. Free, unrestricted and ***open access*** was made available from the INIS Homepage. This provided easy access to reliable nuclear information on the peaceful uses of nuclear science and technology, including nonconventional literature, and made nuclear knowledge readily available worldwide.

Managing diversity at INIS

Managing diversity is a complex process with a number of practical requirements and challenges. Not every element can be fully and successfully addressed, so it is important to take a strategic approach. With INIS, the most important element is international collaboration, followed by the successful diversification of formats, services, and ways to collect bibliographic information and the full-text of related nuclear documents. Demographic diversity is already a given since INIS is part of the United Nations (UN). Everything the UN does is based on two fundamental principles: respect for people and continuous improvement.

The impact of diversity on performance was studied by researchers who identified two impacts: the impact on people and the impact on organizations. According to the cognitive resource diversity theory, which advocates that the cognitive resources of each team member contribute to the overall success of the team, a diversity of cognitive resources promotes creativity and decision making capacity (Simons & Rowland, 2011).

Another group of authors (Cox & Blake, 1991) proposed that cultural diversity impacts six direct aspects of organizational effectiveness, including the capability to attract human resources, cost, the issues of innovation, creativity, and problem solving, the marketing advantages of a diverse workforce, and organizational flexibility. My long experience working with the United Nations, and particularly with the IAEA and INIS, tells me that both aspects have had a positive impact on the success of INIS.

Empowering staff members is much easier than rewarding them in institutions, such as international organizations, since flexibility for rewards usually ends with non-monetary awards and praise. Empowering is encountered more often, usually by encouraging staff to demonstrate their interests, propose new ideas, innovate, be creative, and feel in charge or fully responsible for a particular task. Appreciation often shows remarkable results.

Flexible work environments have a real impact on morale, giving employees a feeling of being empowered and trusted to manage their own workloads. Research shows that flexible work hours are related to higher organizational commitment and job satisfaction (Scandura & Lankau, 1997).

Training and mentoring practices in diversity management should continue, to ensure that there is continuous learning and adaptation. Diversified learning that encompasses technical, as well as soft skills, is important for overall success and special attention is given to that, too. It should be mentioned that besides training its own staff members, INIS regularly offers training to its Member States in the form of a training seminar or in the form of ad hoc and e-learning.

Communication is a never ending story. It includes internal, as well as external means and targets. There are many different ways of communicating today, and INIS is using a good variety of them.

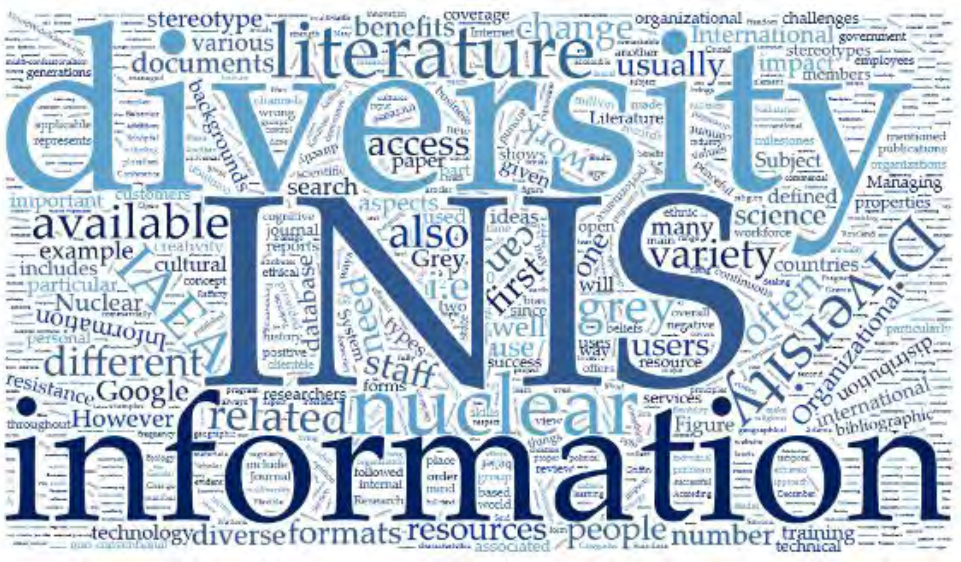
Conclusions

Diversity, defined as the state of having many different forms, types, ideas or properties, is most often associated with cultural diversity or a variety of ethnic backgrounds. However, it is also applicable to information management, usually through the diversity of information resource formats and the variety of information users (i.e. customers or clientèle). INIS represents a very good example of this and its long history of success is proof of that.

Diversity in the work place, and in information management in particular, should be regarded as positive. There is no need to fear diversity. Rather, it should be embraced and used to our benefit. To obtain maximum benefits from diversity, we need to invest time and effort to get to know our customer's needs and to diversify our information base and services. In other words, we need to leverage all available information resources.

It is also evident that diversity, in itself, does not directly bring about benefits. We need to engage in getting those benefits and we need to promote information services, the use of available information resources, and increase visibility and accessibility, particularly of grey literature, as it is not the type familiar to all users. This could be done through continuous training, education and persistent efforts to increase information literacy.

In summary, proper use of diversity in information management can increase the use of available information resources and therefore, impact creativity and improve overall productivity and staff attitudes.



Bibliography

Cox, T. H., & Blake, S. (1991, August). Managing Cultural Diversity: Implications for Organizational Competitiveness. *The Executive, 5*(3), 45-56.

Grama, B., & Todericiu, R. (2016). Change, resistance to change and organizational cynicism. *Studies in Business and Economics, 3*(11), 47-54.

IAEA. (2010). *Subject Categories and Scope Descriptions.* Vienna: INIS.

Kanahara, S. (2006). A Review of the Definitions of Stereotype and a Proposal for a Progressional Model. *Individual Differences Research, 4*(5), 306-321.

Ng, T. W., & Feldman, D. C. (2012, November). Evaluating Six Common Stereotypes about Older Workers with Meta-Analytical Data. *Personnel Psychology, 65*(4).

Rafferty, A., & Griffin, M. (2008). Organizational Change. In *The Sage Handbook of Organizational Behavior. Vol. 1: Micro Approaches* (pp. 603-620). London: Sage Publications.

Scandura, T. A., & Lankau, M. J. (1997, July). Relationships of Gender, Family Responsibility and Flexible Work Hours to Organizational Commitment and Job Satisfaction. *Journal of Organizational Behavior, 18*(4), 377-391.

Schöpfel, J. (2010). Towards a Prague Definition of Grey Literature. *Twelfth International Conference on Grey Literature, December 6-7, 2010.* Prague: TextRelease.

Seid, G. M., & Venkataram, B. M. (2016). The relationship between workplace ethics and employee attraction, loyalty, and financial performance. *International Journal of Applied Research, 2*(5), 585-596.

Shultz, K. S., & Adams, G. A. (2007). *Aging and Work in the 21st Century.* New Jersey: Lawrence Erlabaum Associates, Publishers.

Simons, S. M., & Rowland, K. N. (2011). Diversity and its Impact on Organizational Performance: The Influence of Diversity Constructions on Expectations and Outcomes. *Journal of Technology, Management and Innovation, 6*(3).


International Nuclear Information System
INIS
organizing the world's information on nuclear science and technology and making it universally accessible for peaceful uses
over 150 Member States and international organizations
millions of citations and abstracts published worldwide
hundreds of thousands of full text non-conventional 'grey' literature
multilingual thesaurus in Arabic, Chinese, English, French, German, Japanese, Russian, Spanish
www.iaea.org/inis
IAEA
International Atomic Energy Agency

Looking for Information that is not Easy to Find: An Inventory of LibGuides in Canadian Post-Secondary Institutions Devoted to Grey Literature

Marcus Vaska, Knowledge Resource Service (KRS), Alberta Health Services (AHS)
Rosvita Vaska, University of Calgary (retired), Canada

In today's academic society, one of the most popular web-publishing mediums is a university or college LibGuide. Succinctly defined as a content management and publishing system (Giustini, 2016), these sources of information promote knowledge and learning, and are often seen as the perfect solution for busy librarians and subject specialists. Founded in 2007 by Slaven Zivkovic (Springshare, 2013), an entrepreneur with several years' experience working in libraries, and currently hosted by Springshare, a library technology vendor, LibGuides are presently used by 65,000 librarians and subject specialists across more than 5,000 libraries in 78 countries (Springshare, 2016). A key feature of LibGuides is that anyone can create, use, and learn from them (Giustini, 2016); as a collaborative venture, these guides not only enrich one's knowledge, they also connect with information creators and seekers worldwide. Further, information produced and available on LibGuides is often material that may not be seen anywhere else; this unique content thus supplements any research pursuit.

Libraries use LibGuides to create curricula on a specific subject for a course, promote library collections, and share information within one's faculty, either locally, nationally, or internationally. As an information portal, content within LibGuides can be seen as a living document, constantly being updated, changed, or enriched. Via the use of a WYSIWYG (What You See Is What You Get) content creation approach, users can select templates, copy from existing LibGuides or start from a blank page (Giustini, 2016). Due to its flexibility and adaptability, virtually any LibGuide available on the web can function as a template for a new or existing subject guide. Various pages are linked together via the creation of tabs, functioning in much the same way as the tabs present in an Internet browser. Further, due to the web 2.0 nature of LibGuides, users are able to integrate multimedia content, such as embedding YouTube videos, Twitter feeds, or Chat widgets directly into their pages (Hamilton, 2010). As a collaborative tool, LibGuides can be set up to allow co-owners, even multiple users, to edit content within a guide.

One cannot stress enough that grey literature is a vital source of unique information, often far more current than commercially published material, and circulating at a much faster pace than conventional academic journals. Further, due to its almost unrestricted character, grey literature helps to prevent bias, opening doors to new and emerging research. We believe that LibGuides are a type of grey literature and have a valid place among grey literature supporters, believers, and researchers. In our view, LibGuides should be considered for inclusion into the grey literature typology (GreyNet, 2016), as they are published online, easily accessible, provide up-to-date information, promote self-learning, and contain information that is often overlooked and neglected.

In our paper, we aim to investigate grey literature LibGuides within university and college communities across Canada, documenting categories of areas and disciplines, showcasing how grey literature LibGuides play a pivotal role in research pursuits at academic institutions.

LibGuides: A Brief History

Whether searching through a card catalogue or "flipping" through pages on an iPad, research aids fulfill their role in improving the discoverability of information. This quest for information has resulted in more and more educational institutions resorting to LibGuides, "a simple-to-use, cloud-based web-content management system…to improve the patron experience in finding information" (Dobbs et al., 2013). While the current definition of LibGuides may lead one to believe that this information-rich resource has only existed over the past few years, as a result of enhanced Internet and content creation possibilities, precursors to LibGuides existed decades earlier, most notably print bibliographical

pamphlets in the late 19th century, recommended reading booklets from the 1950s (Dalton & Pan, 2014), culminating in print pathfinders, which appeared on the library scene in the early 1970s (Dobbs et al., 2013). Attributed to Eloise Harbeson, a librarian from Florida State University, pathfinders were considered to be research aids and checklists for seeking out diverse sources of information on a particular topic. Harbeson's pathfinder adhered to a 12-step process, including discussion of the topic on which the guide was created, description of the scope, relevant subject headings, and perhaps most importantly, "emphasis on the process of searching." (Dobbs et al., 2013) Today, these pathfinders or LibGuides are often created locally by the subject librarian, with the purpose of increasing information literacy of the clients they serve. (Bowen, 2014)

In the library world, we often speak of the need to collaborate, to develop connections between the information seeker and the information provider. Regardless of the format that a LibGuide may take, the ultimate premise or goal remains universal: "to connect users with library resources" (Dobbs et al., 2013). Thus, emphasis is placed on library staff to create effective learning content and objectives that are desired by the user and most relevant to his/her information seeking needs (Bowen, 2014).

The current rendition of LibGuides is associated with Springshare, the dominant hosting platform for LibGuides that was launched in 2007 by Slaven Zivkovic, an ed.-tech entrepreneur (Springshare, 2013). With no limit on the number of subject librarians that can use LibGuides, SpringShare closely adheres to its mandate of developing "engaging web applications for libraries and educational institutions" (Springshare, 2013). Despite the endorsement for LibGuides by librarians such as Buffy Hamilton, along with more than 5,000 libraries worldwide, the transition from pathfinders and other precursors has not gone smoothly for all. In 2013, a Google Group was formed to discuss the trials and tribulations of this platform, with users "encountering jargon, inconsistent language, and visual clutter" (Giustini, 2016). As a result, usability testing along with a style guide was implemented. Further, with a notable distinction of being "the primary proprietary guide-creation platform within the library world" (Bowen, 2014), LibGuides permit libraries to customize and alter the standard template with their institutional logos, motto, and other branding methods.

Time is a precious commodity; improving "patrons' experience in finding information" (Dobbs et al., 2013) is a trait that cannot be overlooked if LibGuides are to be successful. With so many decisions regarding the acquisition of library resources dependent on budgetary constraints, presenting evidence-based cases documenting reasons behind a LibGuide's value must be considered at all times. Dean Giustini, a biomedical librarian at the University of British Columbia in Canada, has created a page specific to SpringShare LibGuides on his international HLWIKI (Giustini, 2016). LibGuides are able to detect when a user is accessing a guide on a mobile device, and thus present the content in a refined way that makes it easily accessible and viewable on a smaller screen (Mokia & Rolen, 2012). The continuous need to evolve and present "the latest incarnations of the traditional subject or help guide" (Dalton & Pan, 2014) led to the release of LibGuides v2 in 2014. This new version was recently adopted by both the Knowledge Resource Service and the University of Calgary.

LibGuides as a Grey Literature Document Type

Since launching in 2004, GreyNet, the international grey literature network service has identified 150 document types that are considered to fall within the realm of grey literature (GreyNet, 2016). The latest document type, LibGuides, was added to the GreyNet repository in May 2016. As a living document, a LibGuide is continuously being updated, edited, and enriched, whether it is the addition of a new link, embedding a video, or removing an expired RSS feed. As has often been mentioned in the grey literature community and elsewhere, enhancements in technology, namely the rise of the Internet, has been a direct contributor in unveiling grey resources that were previously hidden. As with grey literature, LibGuides "seamlessly push [students] to other streams of information, including blogs, databases, and other essential library resources" (Hamilton, 2010).

At both the University of Calgary and the Knowledge Resource Service (KRS), it can be difficult to convince the overworked researcher or clinician of the importance of considering grey literature in a research endeavour. Leading by example, however, may help to get the point across. A number of institutions, including the Health Sciences Library System (HSLS) at the Universty of Pittsburgh have implemented courses on responsible literature searching, integrating grey literature as supplementary to mainstream database retrieval (Gerberi et al., 2012). This strongly supports the inclusion of a grey literature section in a LibGuide devoted to material that goes beyond the core databases.

Ever since grey literature gained acceptance as a wealth of resources that supplements the traditional database search, numerous arguments have arisen highlighting the impact and value of the grey. Along with features such as innovative, unique, current content, rapid publication, and freely accessible via an Internet connection, it is no wonder that many researchers see this material as complimentary, "fill[ing] gaps of traditional publishers" (Giustini, 2016). In addition, enhanced web 2.0 features, mobile-optimized content, direct integration with social media outlets such as Facebook and Twitter, customizable interfaces, and perhaps most importantly, excellent customer service, LibGuides brings together "a vibrant user community and a dedicated community website…to share ideas, tips, best practices [and] feature requests" (Giustini, 2016). Table 1 describes core features akin to both grey literature resources and thus, by their nature as a grey literature document type, LibGuides as well.

Feature	Grey Literature	LibGuides	Comments
Innovative, unique content	√	√	Ease and flexibility of creating content to serve a particular nuance in time of need
Rapid publication	√	√	Peer-review process for academic journals can be extremely time-consuming
Currency	√	√	Information available almost instantaneously once it is created
Freely available	√	√	Free, open, immediate access in most cases
Unconventional formats (blogs and other social media)	√	√	Register and receive e-mail updates as soon as new content on a top is available

Table 1: Features pertaining to grey literature and LibGuides

When the 2009 H1N1 pandemic occurred, an alternate use for Twitter, at that time a relatively new social media tool, was discovered. News regarding outbreaks and vaccination locations were instantaneously shared amongst the community at much faster rates than could be reported by news agencies. This fact supports the belief that the emphasis on the inclusion of grey literature is dependent on the discipline. The field of health sciences, perhaps more than most other subject areas, is heavily reliant on traditional publishing methods to secure research grants and buy-in with a proposed idea or methodology (Royal Roads University, 2016). An interesting anomaly when one considers how valuable a non-traditional means of information dissemination was with the H1N1 case study described above.

One of the challenges experienced by many who consult the grey literature is the lack of bibliographic control, thus potentially hindering the discoverability of this material. LibGuides rise above this challenge by allowing the creator to implement and organize categories in LibGuides with virtually limitless possibilities. This ease of usage is a direct correlation to users returning time and time again. To keep track of unique visitors, LibGuides allow the creator to generate statistical reports to see the immediate effect that a guide has on its users and why some sections of a guide are consulted more frequently than others. Many institutional repositories in place today, in and of itself collections of grey literature, allow content contributors to view the number of times their thesis has been viewed and from where.

Many libraries, in both academia and the government sector, are trending towards more of a virtual presence in lieu of a physical location. With communication occurring electronically rather than in person, a sound collaborative structure is needed in order to accomplish what is required in a short amount of time. Just as certain grey literature document types, such as blogs, may involve multiple editors, the same holds true for LibGuides. This enables the guide to maintain a dynamic presence on the web instead of a static document that is soon discarded. A well-maintained LibGuide can, in the words of Buffy Hamilton, "improve the quality of library instruction and service" (Hamilton, 2010).

When pathfinder subject guides were introduced in the 1970s, they were never intended to be comprehensive, all-encompassing repositories containing all of the information available on a topic. Further, they were not intended for experienced researchers who were already well-versed in finding information in their subject area by themselves. In reference to a document on library pathfinders written in 1973, Dalton and Pan agree with this assessment, viewing these guides as intended "for beginners who seek instruction in gathering the fundamental literature of a field new to them in every respect" (Dalton & Pan, 2014). In many ways, this same methodology is used when introducing grey literature. Baring a few organizations, such as the Canadian Agency of Drugs and Technologies in Health (CADTH), which has produced an exhaustive checklist of grey literature resources in virtually all aspects of the health sector, many instances of grey literature inclusion in LibGuides, in particular those seen in the environmental scan of LibGuide use across Canadian post-secondary institutions, reveals a trend to focus on core grey literature resources only, letting the inquisitive mind explore beyond if desired.

How Libraries use LibGuides

According to Springshare, there are presently more than 400,000 LibGuides in existence; this number continues to grow on an almost daily basis (Springshare, 2016). While health sciences disciplines may seem as a logical first choice in which to develop a LibGuide, usability studies have shown that several subject areas, including business, education, humanities, engineering, and social sciences, all contain LibGuides in their respective fields (Quintel, 2016). Due to the ease of creation and maintenance, many libraries have expanded the original subject guide purpose to include courses, e-book collections, documentation on the research process, raising awareness of new technologies, training platforms, and library home pages (Dobbs et al., 2013). With the possibility of creation in a manner of minutes, and minimal requirements for maintenance and updates, LibGuides have also been used to trend worldwide current events (Dobbs et al., 2013).

Be it from students, faculty, investors, or the community at large, receiving acceptance and buy-in from all parties is essential in determining a LibGuide's success. As a case in point, staff working within the Knowledge Resource Service (KRS) department in Alberta Health Services (AHS), are cognizant of time constraints and geographic barriers that may prevent in-person learning. As a virtual learning environment providing library services to all AHS staff in the province of Alberta, KRS staff advocate a flexible blended learning approach, promoting "use of multiple instructional formats such as face-to-face and online activities" (Dobbs et al., 2013). This flexibility, following in the footsteps of an approach vindicated by Dobbs, Sittler, and Cook (Dobbs et al., 2013) is in essence an adaptation of the American Society for Training and Development, offering the best of both worlds, presenting opportunities and challenges associated with in-person and online training.

Once a LibGuide has been launched, it can function as a powerful teaching tool. In order to maintain a stronghold among numerous library resources competing for a client's attention, LibGuides must be regularly updated. Baker (2014) cautions against aggressive and excessive addition of content (the 'kitchen sink' approach); despite one's best intentions, too much content "results in 'cognitive overload" (Baker, 2014).

In a seminal article discussing the implementation of LibGuides at an academic institution, Duncan et al. (2015) introduce four best practices that should be present in all LibGuides: labels and language, layout and uniformity, website integration, and usability. While LibGuides is the dominant term used to describe these information resources used amongst libraries, Duncan and colleagues are quick to point out that this is actually the brand name of the Springshare company; should the SpringShare company change

ownership, the LibGuides name would likely become defunct. This distinction is essential and had led to numerous synonyms presently in use at various institutions to describe these information repositories. For instance, KRS has settled on subject guides, while the University of Calgary uses research guides, terminology favoured by 77% of respondents (Duncan et al., 2015). Further, Duncan et al.'s study gives considerable mention to the importance of consistently labeling tabs, menu items, and headings in LibGuides from the viewpoint of the user rather than the librarian. This is undoubtedly an important recommendation, however, just as importantly, "research guides should use clear, consistent language that is meaningful to the target audience" (Duncan et al., 2015). Since LibGuides are or should be equally represented in the humanities, it would bode well for a discipline advocating the learning of a second language to present a LibGuide in that language of study.

Both KRS and the University of Calgary recently overhauled the layout of their LibGuides to meet user needs and maintain consistency in the content presented. This principle follows Duncan et al's (2015) best practice of layout and uniformity. Minimizing scrolling, documenting the scope and purpose of a guide, providing contact information of either the guide creator or the institution are key, adhering to the recommendation that "guides should be designed to meet patrons' information needs" (Duncan et al., 2015). Focusing on core content within a page is more important than including all of the information that may be available.

Prior to embarking on subject guide creation within KRS, a handout was produced to guide all aspects of implementation, including content requirements and liaising with subject matter experts. Duncan et al. (2015) share this sentiment by emphasizing the necessity of a template guide regarding navigation, standardization, and content management. This iterative process ensures that despite some variation that may be present across LibGuides, the overall website design remains consistent. This notion of consistency bodes well with integrating the guides into curriculum content to ensure they are easily accessible and regularly used. As a case in point, staff at both KRS and the University of Calgary are cognizant of the need to make the LibGuides mobile friendly (since many may be accessing the guides from a mobile device), and possibly explore the presentation of core content from the guide via a mobile app.

Usability, the final best practice advocated by Duncan and colleagues (2015) stresses the importance of feedback regarding navigation experience of LibGuides directly by the users who consult them. A user-centered design and usability testing is important throughout all "stages of the design process" (Duncan et al., 2015). Regardless of the approach taken, perhaps the primary take-home point is communication and consultation between guide owners, subject experts (liaisons) and the institution.

Danielle Becker, a seasoned law librarian in the United States presents LibGuides as entities that rebuild existing content from a website rather than a separate standalone product. While some of her reasoning surrounding the choice of LibGuides may seem self-explanatory, in particular the user-friendly design, social media integration, and elimination of downloads that the SpringShare platform provides, the focus of her case study is on optimizing LibGuides to suit a mobile device environment. (Becker, 2014) Since we live in an ever-connected society, Becker recognizes that an increasing number of tasks are being completed on mobile devices while commuting to work or school (Becker, 2014). While Becker acknowledges the limitations of content on mobile devices, namely the visual appeal and scrolling requirements, with a few modifications, including streamlining content and, above all, usability testing, LibGuides can effectively "repurpose…existing website for acceptable performance on mobile devices" (Becker, 2014).

As perhaps any librarian or subject specialist working in academia can attest to, it can be difficult to maintain a regular in-person presence in a course beyond the 'one-shot instruction lesson. As library instruction sessions are often scheduled at the beginning of term, the skills presented to students are often not put to use until considerably later, when the deadline for a research assignment is looming. Therefore, LibGuides' strength lies in the time-of-need learning approach, guiding "students through a specific assignment" (Baker, 2014) as stand-alone tutorials. Thus, while the librarian is not physically present when the

student consults the guide, he/she maintains a virtual presence by the means of which the information required is presented.

Without downplaying the technical expertise that some librarians and information professionals may indeed possess, Dalton and Pan (2014) also acknowledge the fact that librarian's skills are best utilized when they share their knowledge and expertise with their users by creating, organizing, and disseminating content. They should not be hindered in this endeavour by overtly complicated information technology (IT) procedures. As Dalton and Pan further explain, lack of programming skills is not a hindrance in LibGuides, as even those with limited IT skills will be able to create rudimentary guides pertaining to their subject fields. Further, the library world today often emphasizes the need to collaborate and share one's knowledge with the community. LibGuides content is often re-used within guides and even across guides, adhering closely to the old adage, why reinvent the wheel? Although not necessarily an intent of LibGuides, Dalton and Pan recognize that self-mediated and self-directed learning is a new trend. The multidisciplinary nature of today's programs (shying away from traditional structured formats) has led to increased personalization and individualization. LibGuides embraces this philosophy, supporting self-directed learning and stressing that "the way in which content is presented, not just the volume of information, can determine if it is manageable or overwhelming" (Dalton & Pan, 2014)

In their paper on LibGuides for international students, Han and Hall (2012) describe the creation and maintenance of LibGuides via a demographic group that may seem to be overlooked, but which should never be forgotten: international students. The authors thus present the advantages of LibGuides with this diverse clientele group in mind. First, the consistent user interface presented across LibGuides allows for a high degree of familiarity. As such, Han and Hall mention how several international students studying in the United States have continued to use LibGuides in their home countries (China, Canada, South Korea, Japan, and others) upon their return. Second, the LibGuides interface is diverse and ever-changing, far exceeding the capabilities of a stand-alone webpage. Third, LibGuides have been optimized for display on mobile devices. As Han and Hall go on to explain, these feature is of tremendous benefit to international students, who often access content (both academic and social media) on the go from their iPad or iPhone. Finally, many students, whether or not they study abroad, may never have the opportunity to meet their subject librarian in person; the majority of communication is thus done in a virtual setting (chat, Lync, Skype, etc.) LibGuide's interactive features, including a photo of the librarian along with his/her contact details ensures students will continue to receive support for their research needs wherever they may be.

The Investigation: Grey Literature LibGuides within Canadian Universities and Colleges

Despite the ease within one can create, adapt, and reuse instructional material in LibGuides (Bowen, 2014), few studies exist on the benefits of this information commodity to the intended target audience, namely students. In order to obtain a representative sample of the use of grey literature in LibGuides across Canadian post-secondary institutions, we conducted an environmental scan, identifying 17 colleges or universities where grey literature resources were directly mentioned and included alongside academic databases (Table 2)

College/University	City	Province/Territory
Bow Valley College	Calgary	Alberta
University of Alberta	Edmonton	Alberta
University of Calgary	Calgary	Alberta
Simon Fraser University	Burnaby	British Columbia
University of British Columbia	Vancouver	British Columbia
University of Victoria	Victoria	British Columbia
University of Manitoba	Winnipeg	Manitoba
Dalhousie University	Halifax	Nova Scotia
Carleton University	Ottawa	Ontario
Lakehead University	Thunder Bay	Ontario
University of Toronto	Toronto	Ontario
University of Waterloo	Waterloo	Ontario

University of Western Ontario	London	Ontario
York University	Toronto	Ontario
University of Prince Edward Island	Charlottetown	Prince Edward Island
McGill University	Montreal	Quebec
University of Saskatchewan	Saskatoon	Saskatchewan

Table 2: Canadian Post-Secondary Institutions Containing Grey Literature Resources in LibGuides

After viewing the LibGuides within each of the post-secondary institutions listed in Table 2, 52 library staff (librarians and information specialists) were identified. A brief online survey was sent to each of the 52 library staff members, to uncover how students and researchers use grey literature, and perhaps most importantly, to verify from the participant responses whether or not sections of existing LibGuides have been devoted to including the grey literature in information-seeking pursuits (Table 3).

1. Is grey literature mentioned in subject guides/research guides/LibGuides within your institution?
2. If your response to question #1 is yes, please list the top 3 grey literature resources that are mentioned in the subject guide(s)/research guide(s)/LibGuide(s).
3. Do you encourage your students to consult grey literature resources in their research? (if not, please explain why not)
4. What is your discipline/area of study?

Table 3: Use of Grey Literature in Canadian Post-Secondary LibGuides: Survey Questions Results:

Nine of the 17 institutions polled participated in the survey, yielding a response rate of 52.9%. All respondents confirmed that grey literature was mentioned in the research guides/subject guides/LibGuides used within their institution. Further, all respondents strongly believed that it was important to encourage students to consult grey literature in their research, particularly if conducting scoping or systematic reviews. With the rigorous methodology required in these publications, the prevalence of grey literature and its stance as a supplement to searching in the databases cannot be overlooked. In a recent seminal paper quantifying grey literature search trends in systematic reviews, Saleh et al. (2014) indicate that grey literature comprises 27% (an average of 6.5) hours of total searching time across all resources. While this figure is specific to research conducted in the health sciences, one survey respondent explained that grey literature plays a key role in both Masters of Education and PhD in Education programs.

One respondent emphasized the importance of grey literature in the Canadian context, especially with regards to public policy questions. As we can attest to, there is no true mainstream research database used in academia focusing only on Canadian publications; the majority of articles retrieved from a literature search are published in the United States, and thus the need to supplement these resources with Canadian content wherever possible is increasingly important.

When questioned why librarians and information specialists encourage the use of grey literature in research pursuits, one participant elaborated in detail on the numerous benefits of introducing grey literature in a search strategy, including currency of information, rapid publication, contains unpublished research, is produced by experts in the field, and has a wide geographic scope, from municipality to province and beyond.

Beginning in the early 1990s, grey literature was cited in the “white literature” (i.e. medical journals), consisting mainly of technical reports that provided reliable data on research in progress. (Alberani, De Castro Pietrangeli, & Mazza, 1990) It thus came as no surprise to the authors of this paper, following both the environmental scan and online survey, that the majority of respondents maintained LibGuides within various domains of the health sciences. Nevertheless, three respondents were adamant about the value of grey literature in their subject specialties, including education, international affairs, and political science.

When asked to list the top three grey literature resources most frequently mentioned in their research guides/subject guides/LibGuides, we were pleased that 1/3 of respondents mentioned either GreyNet, OpenGrey or the New York Academy of Medicine's monthly grey literature reports. Further, it was interesting to see many of the 150 grey literature document types identified by GreyNet represented (Table 4).

Agency for Healthcare Research and Quality (AHRQ)
Canadian Council on Learning
Canadian Government Documents
Canadian Health Research Collection
Canadian Public Policy Collection
Columbia International Affairs Online
DocuTicker
eLibrary: Conference Board of Canada
ERIC Education Documents
Grey Literature Report (New York Academy of Medicine)
Grey Matters: Canadian Agency for Drugs and Technologies in Health (CADTH)
GreyNet
Health Technology Assessment (HTA)
Institute of Health Economics
Ontario Public Health Libraries Association (OPHLA)
OpenGrey
Registry of Open Access Repositories (ROAR)

Table 4: Different Types of Grey Literature Resources Identified

Ideas on Promoting Grey Literature in LibGuides

With a switch in philosophy from library-mediated to self-mediated, the success achieved by LibGuides lends to its core benefits, namely saving "the time of users by allowing them to assist themselves", and saving "the time of librarians by allowing users to answer their own commonly asked questions" (Dobbs et al., 2013). This is particularly relevant in a distance education environment, where learners are located in different corners across the globe. Rather than specify that the user must come to the library, the library waits for the user; with a reliable Internet connection, students can access content within LibGuides at their convenience, day or night, "without any librarian mediation" (Dobbs et al., 2013).

While there are undoubtedly many different methods available to convey information in lieu of a LibGuide, these one-stop information shopping centres have grown exponentially since their creation less than a decade ago. With the ability to customize particular sections more easily and with less required technical knowledge than most web pages, LibGuides reduce "technical and temporal barriers" (Bowen, 2014) to producing and publishing information as quickly as possible, an endearing value.

While it may seem self-explanatory to many, Denise FitzGerald Quintel, a web applications librarian at the University of Alabama libraries, stresses the importance of usability studies to determine if a LibGuide is benefitting the intended audience and is actually what a user wants. According to the literature, LibGuides may be used as an instructional aid to introduce curriculum concepts, they are above all, tools "means for independent learning and resource discovery" (Quintel, 2016). Quintel further elaborates that LibGuides should be viewed not only as a quantitative measure to present evidence-based information to users, but also as a qualitative exercise to determine a user's gut reaction to a guide: is there awareness that the guide exists, is it useful, and will it help achieve a desired goal? In her study, Quintel found that 63% of students had no knowledge that a LibGuide existed for their course until they attended a library instruction session.

Quintel concludes her usability case study with several suggestions that anyone considering a LibGuide for his/her institution would do well to adhere to. Within each of the 6 recommendations presented, the underlying factor is to cater the guide to the user's needs and expectations, not to the content creator's. The long-held adage that the customer is always right bodes well for the creation of a LibGuide, and for the inclusion of

grey literature content within that LibGuide. Perhaps Quintel says it best with her proclamation that "the beauty of usability testing is that with small and iterative changes, we have a certain freedom to enact solutions" (Quintel, 2016).

While there are countless benefits supporting in-person library instruction and training, students, researchers, and healthcare professionals residing in remote rural areas may never have the opportunity to "meet" beyond a computer screen. Librarians must thus put themselves into the mindset of their learners, questioning what engages them to learn, and modifying the content within their LibGuides to address this need. Collaborating with stakeholders, be it faculty, students, and other professionals is key, a requirement exemplified by Baker: "what we need to do as educators is develop a way to give students the information skills they need, when they need them, and in a way that fits their learning styles and preferences" (Baker, 2014).

Looking into the Future

According to Bowen (2014), LibGuides are a unique entity, and "no other proprietary products have entered the market" to compete with these online information repositories. Despite not even existing in their current form a decade ago, LibGuides are rapidly replacing static library webpages. As illustrated in Bowen's case studies, ease of use and a relatively low learning curve are essential in ensuring sustainability of LibGuides for years to come. While mention is often made of the busy and never-ending workload of health professionals, and for that matter, researchers in all disciplines, the same can be said for library personnel. Producing and maintaining webpages takes a considerable amount of work; "Libguides greatly reduces the technical and temporal barriers to completing that work" (Bowen, 2014).

In their article advocating LibGuides as a content management system for busy librarians, Verbit and Kline (Verbit & Kline, 2011) advocate for content management systems created specifically for libraries, since these resources need to be flexible to consider the far-reaching visions at libraries within academic institutions. While LibGuides does not try to promote itself as the be-all-to-end-all only solution for organizing library resources in a logical manner, the ability to reuse and easily transfer content across multiple guides, coupled with the integration of statistics, images and videos, catalogue records, and optimization for mobile devices are highly touted benefits that cannot be ignored. Looking into the future of library services, librarians considering the implementation of LibGuides as a means of advertising the library resources available at their institutions would do well to adhere to LibGuides' strengths "for information literacy instruction and just-in-time online learning modules" (Verbit & Kline, 2011).

While LibGuides have been seen in a positive light whenever they are introduced as the new primary content management system to showcase a library's resources, the importance of pilot testing and evaluation should never be taken lightly. Lack of bibliographic control is one of the primary reasons that grey literature can be difficult to locate; as such a grey literature search can take considerable longer to conduct than running a search through an academic database, which due to the search filters working in the background, is more streamlined and efficient. Dalton and Pam (2014) acknowledge this challenging and time-consuming aspect, yet feel it can be overcome by streamlining a subject guide in much the same manner as a database. As a case in point, we are often astounded that Google continues to be the go-to information source for students on a regular basis, despite the thousands of dollars invested each year into ensuring that high-quality evidence-based information is available at one's fingertips. Perhaps this is why many LibGuide creators have increasingly adapted a "one stop shopping" approach, a theme that is likely to be particularly useful for students, who "often report feeling overwhelmed by the volume of resources and information available" (Dalton & Pan, 2014).

Due to the considerable time commitment that LibGuides require, they are seen as living documents that will exist for years to come, constantly adapting according to client requests and needs. With this comes a strategic management plan, ensuring that LibGuides continue to be managed and updated on a regular basis. This maintenance correlates well with the next stage of implementation, namely promotion and marketing. A hidden guide or one that no one is aware exists will not generate the traffic required to justify its continued maintenance and use. Part of this may be confusion with what is meant by the term

LibGuides, as according to the literature, it is often used interchangeably with subject guides, research guides, and general guides. Further, social media is accepted as a legitimate means of creating awareness, and thus Dalton and Pan (2014) advocate for numerous channels at one's disposal to ensure the guide will become a known entity. In addition to social media, e-mail, printed flyers, and word of mouth (as well as other forms of informational communication) should all be considered.

Finally, the pilot LibGuide must be evaluated, not only via quantitative means (usage statistics), but rather qualitatively as well, ensuring that the vices of all potential users of the LibGuide (students, faculty, staff) are represented and receive the opportunity to share their feedback and provide comments. While LibGuides can be created and presented as independent identities with no institutional affiliation, they are seen as most effective "when embedded within an overall support framework and strategy" (Dalton & Pan, 2014). A benefit of consulting these resources is that they often contain society publications or author manuscripts that are available almost instantaneously.

In their paper on the sustainability and future of online research guides, Giullian and Zitser (2015) acknowledge LibGuides' many benefits, including engagement of students in their own learning and marketing by librarians to raise awareness of this online information repository. Despite the valiant efforts of subject librarians, students will, on a consistently regular basis, look up information on a topic in Google before even considering another possible source. This phenomenon of strictly sticking to web-based resources as a search methodology is akin to the "impossible task of drinking from a fire hose" (Giullian & Zitser, 2015). Within a seemingly infinite number of libraries turning to LibGuides, the uptake has seemed almost too quick at times, and there is concern that it has become oversaturated in the past few years. With so many LibGuides arising from a paper version, Giullian and Zitser (2015) question whether home-grown born-digital guides, which due to several competing factors, namely time, server space, and money can be modified to ensure LibGuides sustainability for many generations to come.

One of the primary benefits of LibGuides is the ability to re-use content across various guides within one's home institution. Further, while various aspects of a LibGuide can be customized to a creator's desires, the underlying principle is that LibGuides are essential a rod-sourced community of research guides that can be used as discovery tools by virtually anyone on the planet" (Giullian & Zitser, 2015). This can be either a welcome feature or a potential drawback. With the Open Access movement gaining more and more momentum, creators of LibGuides have jumped on the proverbial bandwagon and articles, theses, and guidelines that can be accessed from anyone around the world via a few clicks of a mouse button. However, this sharing notion does not extend to individual databases or journals. Academic institutions pay hundreds of thousands (even millions) of dollars annually to ensure that their intellectual property is protected and accessible only via an individual affiliated with that institution. With any economic crisis, smaller academic institutions are affected, and perhaps echo the sentiments of Giullian and Zitser in that they "can admire the merchandise on display, marveling at its opulence and availability, but cannot actually afford to use it" (Giullian & Zitser, 2015).

How does the library community sustain LibGudies for many future generations? Both authors of this paper align in their thinking with Giullian and Zitser's thoughts on customizing content specific to a specific assignment or time of need rather than presenting a haphazard collection of seemingly random resources. Further, the importance of engaging the user, whether it is a student who will be using the LibGuide for a few random intervals during a term, or a faculty member that refers to it daily, or even a patient who requires more information on his/her condition, the user has a right to have his/her say in the content being created for LibGuides and voice his/her opinion should he/she disagree with what is being presented. In the brief survey around which this paper is based, respondents, whether explicitly or implicitly, acknowledged the importance of a collaborative relationship between the librarian and teaching faculty at their institution. In Giullian and Zitser's words, this allows "professional librarians to do what they do best: organize, educate, and inspire" (Giullian & Zitser, 2015).

References

Alberani, V., De Castro Pietrangeli, P., & Mazza, A. (1990). The use of grey literature in health sciences: A preliminary survey. *Bulletin of the Medical Library Association, 78*(4), 358-363.

Baker, R. (2014). Designing LibGuides as instructional tools for critical thinking and effective online learning. *Journal of Library & Information Services in Distance Learning, 8*(3-4), 107-117.

Becker, D. (2014). LibGuides remakes. *Computers in Libraries, 34*(5), 18-22.

Bowen, A. (2014). LibGuides and web-based library guides in comparison: Is there a pedagogical advantage? *Journal of Web Librarianship, 8*(2), 147-171.

Dalton, M., & Pan, R. (2014). Snakes or ladders? evaluating a LibGuides pilot at UCD library. *Journal of Academic Librarianship, 40*(5), 515-520.

Dobbs, A., Sittler, R., Cook, D., & Library and Information Technology Association (U.S.). (2013). *Using LibGuides to enhance library services : A LITA guide*. Chicago: American Library Association.

Duncan, V., Lucky, S., & McLean, J. (2015). Implementing LibGuides 2: An academic case study. *Journal of Electronic Resources Librarianship, 27*(4), 248-258.

Gerberi, D., Hawthorne, D., & Larsen, K. (2012). Rethinking responsible literature searching using LibGuides. *Medical Reference Services Quarterly, 31*(4), 355-371.

Giullian, J., & Zitser, E. (2015). Beyond LibGuides: The past, present, and future of online research guides. *Slavic & East European Information Resources, 16*(4), 170-180.

Giustini, D. (2016). *LibGuides from Springshare*. Retrieved September 7, 2016 from http://hlwiki.slais.ubc.ca/index.php/LibGuides_from_Springshare

GreyNet (2016). *Document types in grey literature*. Retrieved September 7, 2016 from http://www.greynet.org/greysourceindex/documenttypes.html

Hamilton, B. (2010). *Top ten reasons I love LibGuides*. Retrieved September 7, 2016 from https://theunquietlibrarian.wordpress.com/2010/04/16/top-ten-reasons-i-love-libguides

Han, N., Hall, S. (2012). Think globally! enhancing the international student experience with LibGuides. *Journal of Electronic Resources Librarianship, 24*(4), 288-297.

Magi, T., & Mardeusz, P. (2013). Why some students continue to value individual, face-to-face research consultations in a technology-rich world. *College & Research Libraries, 7*4(6), 605-618.

Mokia, R., & Rolen, R. (2012). LibGuides: Improving student and faculty access to information literacy. *Codex, 1*(4), 37-45.

Neves, K. & Dooley, S. (2011). Using LibGuides to offer library service to undergraduate medical students based on the case-oriented problem solving curriculum model. *Journal of the Medical Library Association, 99*(1), 94-97.

Quintel, D. (2016). LibGuides and usability: What our users want. *Computer in Libraries, 36*(1), 4-8.

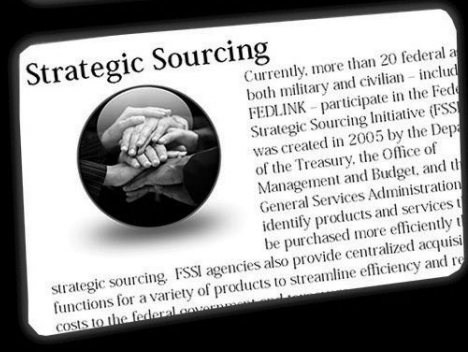
Royal Roads University. (2016). *Grey literature: What is it? the value of grey literature*. Retrieved September 7, 2016 from http://libguides.royalroads.ca/greylit/value

Saleh, A., Ratajeski, M., & Bertolet, M. (2014). Grey literature searching for health sciences systematic reviews: A prospective study of time spent and resources utilized. *Evidence Based Library and Information Practice,* 9(3), 28-50.

Springshare. (2013). About us. Retrieved September 7, 2016 from http://www.springshare.com/about.html

Springshare. (2016). LibGuides by Springshare. Retrieved September 7, 2016 from http://www.springshare.com/libguides/

Verbit, D., & Kline, V. (2011). Libguides: A CMS for busy librarians. *Computers in Libraries, 31*(6), 21-25.



FEDLINK

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

LIBRARY OF CONGRESS
FEDLINK

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

Bridging the Knowledge Gaps: What Grey Literature does for the Library of Congress' Indigenous Law Portal

Carla Davis-Castro, Library of Congress, USA

Abstract
In 2014, the Library of Congress launched the Indigenous Law Portal, an open-access platform with digital resources about how indigenous peoples of the Americas govern themselves. A new classification schedule is being written, *Law of the Indigenous Peoples in the Americas (Class KIA-KIX)*, organizing resources by country and region, legal area, and tribal entity. The classification guides the structure of the online Portal, which includes Canada, the United States, and Mexico to date. The paper will cover the creation of the Portal, the current resources, and the ongoing research on indigenous groups of Central America. Grey literature has been vital to the creation of the name authority records, the classification schedule, and the Portal, whose open access resources serve a diverse global public.

Introduction

In reference to indigenous law, the Law Librarian of Congress' observed that "the subject matter is increasingly important globally but has not received much attention." Dr. Jolande Goldberg is paying attention. In her capacity at the Library of Congress' Policy and Standards Division, she is writing a classification schedule called the *Law of the Indigenous Peoples in the Americas (Class KIA-KIX)*. This schedule organizes resources by nation and region, legal area, and tribal entities. The emphasis on indigenous law is not to set it above or apart from other bodies of law but so it can be made more visible among humanity's laws.

The new classification coincides with broader developments like the *United Nations Declaration on the Rights of Indigenous Peoples* adopted in 2007 and the *American Declaration on the Rights of Indigenous Peoples* adopted June 2016, which outline global and hemispheric indigenous rights, respectively. The challenge for information professionals is to recover and share resources within library collections which constitute the legal record of sovereign tribal governments. For over a decade, the Law Library of Congress has been tracking and analyzing indigenous law sources from its rare book vaults to the far corners of the Web, which along with digitization projects, have increased visibility of indigenous sources.

In response to global trends, user frustration regarding lack of accessibility, and flawed access practices, the Library of Congress (LC) launched the Indigenous Law Portal in 2014 to share digital indigenous law resources. The open access Portal is available on the Law Library of Congress homepage at http://www.loc.gov/law/help/indigenous-law-guide/index.php and links to legal documents, tribal websites, global, national and regional advocacy organizations, and over 400 digitized materials from Law Library of Congress, and general resources. "The huge joy about the Portal is that it brings in cultural, political, and social aspects but also mapping shows how geography affects interpersonal relationships and intrapersonal development" the Law Librarian of Congress said.

The author's paper is enriched by the reflections of 11 colleagues who work or have worked on the Indigenous Law Portal. Quotations from the interviews are shared throughout the paper. As a disclaimer, while interviewees expressed their professional opinions, they may also be personal opinions and do not reflect the official policies of the Library of Congress. As the Portal is a work in progress and the cause of some debate, interviewees spoke freely because their names would not be included with their quotations.

Finding Grey Literature

Political trends complicate the availability and discovery of indigenous legal material and at the same time make the Indigenous Law Portal's mission urgent. Much of the work of the Portal has been shouldered by a few permanent employees at LC. Extensive

contributions have been made by retired librarians, three interns, and one short-term employee. One project manager remarked "we always thought there would be partners like universities and tribal organizations." While many have expressed admiration, LC has found no partners for the Portal thus far. This means that the grey literature available on the Portal either comes from the physical collections of LC or what is freely available online.

Finding relevant resources to include on the Portal is circuitous. One researcher asserted "I have found Canadian documents in Germany. We are as close to Indiana Jones as you can get from behind a computer. We are really hunting this information down." With few direct informational paths it is sometimes necessary to find nontraditional resources: student theses, project statements from nonprofit partners, and social media accounts. These diverse grey literature sources contribute heavily to the Portal's resources for well-known and lesser-known tribes alike.

Cultural respect tempers the quest for comprehensive information on the Indigenous Law Portal. Information seeking has limits as "secret and sacred aspects of Indigenous knowledge have no place in the public domain and should remain outside of LIS systems." [1] For example, the researchers have been cautious in the areas of medical and religious law. While librarians are trained to search for information, the people behind the Portal understand that some information is not meant for public dissemination.

Name Authority Records and Grey Literature

Name authority records (NARs) are the building blocks of the classification, which organizes the Indigenous Law Portal content. However, organization is not straight forward because, as one LC legal information specialist noted, "it can be difficult to find and there can sometimes be a variety of names for a tribe so the work on the Name Authority Records is important." Grey literature frequently supports NAR creation for the Portal.

The first major hurdle is that entangled indigenous names affect NAR creation. Extensive research is necessary because in many countries there has been ethnic cleansing, civil war, constitutional or geopolitical boundary changes, which make it hard to track the changing name(s) of an indigenous group. Over time, names have been lost, recovered, borrowed, reconstructed, and filtered through foreign languages. Sometimes it is not easy to distinguish between a place name, a people's name, and a people's language. As one law librarian explains "we have to be careful about superimposing. You may have the community of people and the village with the same name [i.e.] two different jurisdictions running simultaneously." Qualifiers can be used distinguish different entities with the same name. Kesler writes that "many Indigenous people prefer to identify themselves by specific local terms based on family and community location and traditional names" [2] and in accordance with this notion, the name authority records attempt to include modern and historical names, spelling variations, commonly-used and lesser known names. Some names for indigenous peoples are pejorative but are still included in the NAR because it appears in information sources which may contain other valuable evidence. One example from Mexico are the Ngi-iva/ngiba who were called "popoloca" by Mexica invaders, which can be translated as "mute," "not intelligent," or "barbarian" according to one grey literature resource.[3] This unflattering name has stuck for centuries, was adopted by the Spanish colonizers, and is even the title of the resource. In the case of names, it can be difficult to determine when there is enough evidence to justify a NAR for an indigenous group or how to reconcile differing sources.

As name authority work moves through the continent, more puzzles emerge in trying to understand tribal connections. There are myriad examples of displacement like the Seminole and Kickapoo groups of the state of Coahuila in northern Mexico. These particular communities went south in the nineteenth century, breaking off from their relatives who currently reside in the United States (mainly Florida and Oklahoma for the Seminole while the Kickapoo are principally in Texas and Oklahoma). The challenge was ensuring that the NARs reflect the links between these geographically dispersed yet related communities. Diverse grey literature sources played a role in establishing these critical connections.

On a separate note, the type of name authority used for indigenous nations has recently changed. In 2005, LC decided to move tribal entities from corporate entities to autonomous jurisdictions. "This meant that they had a status like nations," according to one

librarian who spent her career at LC. This required changes to existing NARs and has changed the category for the new ones being created for the Portal.

The significance of the name authority work cannot be overstated as over 2,000 records were created for North America alone, literally putting people on the map. A classification specialist remembered "surfing the web together with my Canadian colleagues" in 2009 and relying on a Canadian resource saying, "When I started looking at how I could build a portal, I started looking at AANDC's (Aboriginal Affairs and Northern Development Canada) gateway to thousands of websites. It has been our major authority on names." Looking outside of the United States for professional assistance helped provide a model for the Portal as well as a treasure trove of grey literature.

Diverse resources are leveraged to build the authority files, a specialty of the Library of Congress. One librarian expands on this idea, "In Latin America, projects are based on who's in office; librarians are political appointees. There are very few authority files in the world and they don't exist in the same way in Latin America and are not easily shareable. It is a contribution that we [at LC] make to the community at large." This contribution is understood by other librarians involved in the Portal where one said "Names are important for identity, showing what the federal government calls them, what they call themselves, what they used to call themselves." Naming what has been historically erased and forgotten is meaningful work and as one librarian said, "It gives groups a certain legitimacy, certain recognition." Fortunately, NARs are also flexible as "names can be changed as more research is conducted, as laws change. The beauty of the authority file and the Portal is that you can navigate through those changes. The Portal is an excellent, amazing view into these entities and their development and their position within their communities." This flexibility means that the NARs associated with the Indigenous Law Portal can respond to changes in the world.

While NARs are established one by one, together they can help organize a global body of knowledge. The Portal is valuable, said one law librarian, because "it's going to organize the records of the communities of our nation and if other countries don't use this method that's OK but we have to have a method for our citizens to do research about foreign countries and the native peoples who reside there." While Americans are the majority of the users using the Portal, a librarian reminds us that "our information goes worldwide, there are no borders." Usage statistics, which are discussed later, show that the audience is indeed worldwide.

Classification and Grey Literature

The Indigenous Law Portal as it appears today is not the final product since the classification schedule is expanding southward throughout the hemisphere. A digital resources librarian expressed her appreciation for "how the Library of Congress classification and name authorities provide a strong foundation for such an ambitious and global project." Another librarian emphasized that such a foundation will serve the Portal into the future because "embedding the classification system is something we can build on later." As the Portal is structured according to the classification, the former will mirror the growth of the latter.

While classification has traditionally been applied to books and other media, a researcher excitedly believed "we are bringing classification into the 21st century…it is now being applied to the Internet. As opposed to an inefficient word search generating 20,000-page response, we are organizing this knowledge in a hierarchical fashion." A classification specialist agreed that "just because it's not sitting on the shelf doesn't mean LC doesn't have it digitally or have access to it." The Law Librarian of Congress echoes this sentiment when she said "on so many levels within the Library of Congress, [the Portal] is a model or example for future classification on the web." This application of classification to digital resources could mean easier access to grey literature.

The Portal leverages the classification for access but also as a research tool. As one librarian expressed it "by making indigenous laws accessible online, they are more easily researched, compared, and studied, bringing more recognition and scholarship to these fields." There is research being done in Central America, for example, much of it by an upcoming generation of scholars writing theses. This body of grey literature examines a

plethora of topics: language revitalization, pluricultural and multilingual school curricula, holistic medical centers, the creation of indigenous territories with their own governments, and emerging models of multicultural legal systems. While there are many ways to access these digital materials, the Portal uses class KIA-KIX to guide users to this unique grey literature.

National Bibliographies and Grey Literature

The name authority records and the classification together help shape the digital resources available on the Indigneous Law Portal. A bibliography of digital resources is created in the form of a spreadsheet for each nation state; the author is responsible for the national bibliographies of Central America which is the next region that will be added to the Portal. These national bibliographies are heavily populated by grey literature documents as evidenced by the following table:

National Bibliographies of Indigenous Law			
Country	Grey Lit	Total	% Grey Lit
Mexico	90	97	93%
Guatemala	175	177	99%
El Salvador	14	19	74%
Honduras	31	31	100%
Nicaragua	65	65	100%
ALL	**375**	**389**	96%

As the numbers demonstrate, the Central American section of the Portal depends upon and leverages diverse grey literature sources to share indigenous law with users.

Recognition and Grey Literature

Recognition by a government is not a prerequisite for including an indigenous group on the Indigenous Law Portal. The librarians rely on self-identification as best as can be determined by information sources at their disposal. Grey literature is often the only resource type that fills this niche and is especially valuable if it includes firsthand accounts of indigenous community members.

On the other hand, external recognition comes in many forms. A community may be recognized by their neighbors, the international community, or another institution. According to one librarian the Portal "is giving a certain weight and validity to indigenous peoples' perspectives. Classification is a huge recognition." For some communities, it may be the first time an outside institution acknowledges them. Upon conducting trainings, some responses to the Portal included sentiments such as "this is the first time these peoples have been evidenced" and "with the name of LC behind it." It remains to be seen what such information-based recognition will mean for indigenous communities in the future.

Collections and Grey Literature

One could spend lifetimes discovering the extensive collections of the Library of Congress and it grows every day, including collections of grey literature. While working on the Indigenous Law Portal, a cartographer remarked, "What surprised me was how comprehensive the website is, and the extent of legal resources it provides for these indigenous Americans." The goal of the Portal is to become equally comprehensive for North, Central, and South America.

While the author conducts research for the Portal, she has the opportunity to observe broad patterns and uniquely local cases in indigenous law. Some legal patterns define the Central American region such as councils of elders which have decision making power in many indigenous communities. Within that there are local particulars such as the number of leaders, the length of their terms, the gender ratios and whether they are elected by votes or selected by consensus. The richness of the many existing legal systems cannot be described adequately here but the following examples highlight grey literature digital

resources culled from the growing national bibliographies of Mexico, Guatemala, El Salvador, Honduras, and Nicaragua.

Example: Mexico

Image is a screenshot from the General Law of Indigenous Peoples' Linguistic Rights' *webpage provided by Mexico's National Institute of Indigenous Tongues*

The *Presentation of the General Law of Indigenous Peoples' Linguistic Rights* was published in 2015 by the National Institute of Indigenous Tongues, which uses the acronym INALI in Spanish.[4] The law itself was published in the Mexican national gazette, *Diario Oficial de la Federación*, several years prior in 2003 but INALI provides the added value of interpreted recordings of the law in 65 indigenous languages. Most of the digital resources on the Portal are written sources but this one is unique because it is mainly an oral resource.

The law, of course, is not an indigenous law but instead a national law. Federal or national governments are prodigious publishers of grey literature and the Portal includes many examples because national laws touch indigenous citizens. All of the countries of the Western Hemisphere have instances of legal pluralism. There may be a national government, regional governments for states or departments as they are commonly called in Central America, local or municipal governments in addition to local and/or regional indigenous governments. These overlapping jurisdictions mean that indigenous peoples are affected by various legal systems and while the Portal focuses on indigenous law, the communities live within a larger national context so some resources such as this one are included. Also, the audio format respects the many unwritten indigenous languages within the country by addressing groups in their native languages in addition to the Castilian (Spanish) of the original law.

Example: Guatemala

Image appears on the Mancomunidad de Municipios Garífunas de Honduras (MAMUGAH) blog http://mamugah.blogspot.com/

Student research from institutions of higher learning provides another rich source of grey literature that is found frequently among the national bibliographies. Stacey Caron Castillo Lewis's 2005 social work thesis from the Universidad de San Carlos can be translated as "The impact of migration on Garifuna cultural identity loss." [5]

As the title indicates, the focus is on this Caribbean people descended from Arawak, French, English, Spanish, and West African peoples. Specifically, it includes information about two Garifuna organizations active in Guatemala: the Black Guatemalan Organization (Organización Negra Guatemalteca ONEGUA) and the Organization of Guatemalan Garifuna Women (Organización de Mujeres Garífunas Guatemaltecas ASOMUGAGUA). Written by a self-identified Garinagu (singular form of Garifuna), she interviewed leaders of these indigenous organizations and examined the ways in which they can fight against the detrimental effects of emigration and subsequent cultural loss. The Portal is enriched by such grey literature that shares findings from research conducted on the ground.

Example: El Salvador

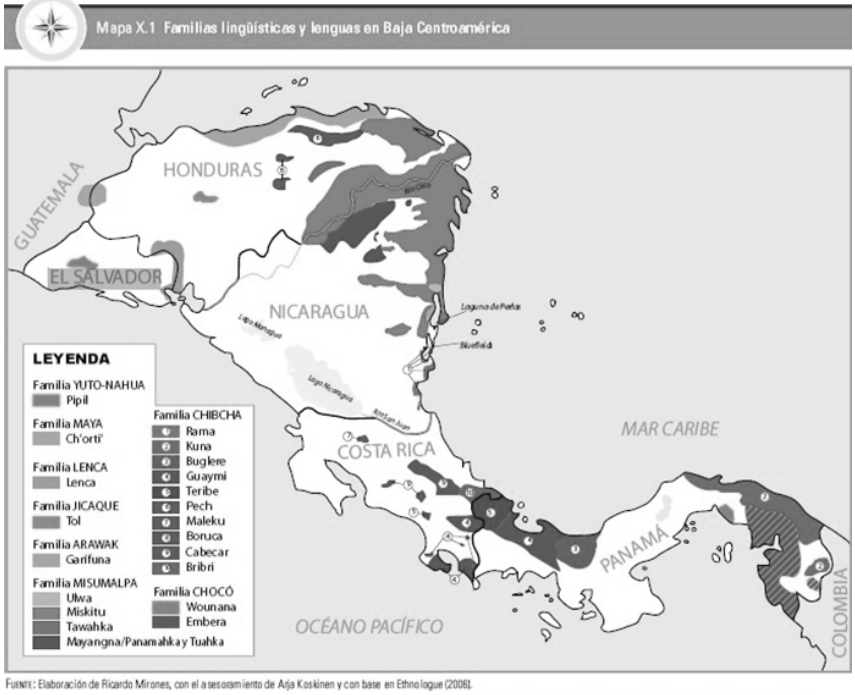


Image shows map of indigenous linguistic families in Central America from UNICEF's Atlas sociolingüístico de Pueblos Indígenas de América Latina, tomo II

Large nonprofit organizations have resources to provide comprehensive, regional studies that allow for comparison and high-level statistics. In 2009 the United Nations International Children's Emergency Fund, better known as UNICEF, and the Foundation for Education in Multilingual and Pluricultural Contexts-Andes published a two-volume work entitled "Sociolinguistic atlas of indigenous peoples of Latin America." [6]

This hefty grey literature source provides important contextual information about dozens of indigenous communities under the heading "research guides," which appears both in the classification and on the Portal. This is an example of a digital resource that does not fall into a national bibliography but rather a regional one. While the Central American national bibliographies are being created, the author is simultaneously collecting global and hemispheric resources or the regions of North America, Central and South America (the term Latin America is not being used), and Central America and South America separately. Again, this digital resource is not explicitly about indigenous law but it does provide contextual information about the Nahua-Pipil, Lenca, and Cacaopera of El Salvador. Furthermore, during the twentieth and twenty-first centuries, many international policymaking bodies, like UNICEF's parent body the United Nations, have created a space for indigenous peoples to express their political will.

Example: Honduras

Image is from Diagnóstico participativo en el pueblo indígena Pech Honduras C.A. *of a meeting of Pech women discussing gender violence in Subirana, no page number provided*

Nonprofit publications are a font of information on indigenous communities. The International Forum on Indigenous Women and the Indigenous Lenca Movement of Honduras published "Participative diagnostics of the indigenous people of Pech, Honduras C.A." where the last letters refer to "Central America." [7] This grey literature resource does not include page numbers or a publication year, which are crucial metadata that the Portal librarians would like to include. Nonetheless, this publication provides valuable insight into the Pech women's understanding of gender violence.

One Pech interviewee referenced a punishment that used to be applied to men who abused their female partners. To paraphrase, some of the leaders believe that the Pech should go back to the traditional punishment of building a fire and tying the man to a wooden frame over the fire to breathe in the smoke. If he could not stay there long enough, he was beaten until he couldn't move.[8] This example of an indigenous punishment for gender violence tells us something about how the Pech responded to this human rights issue as expressed by their law.

Example: Nicaragua

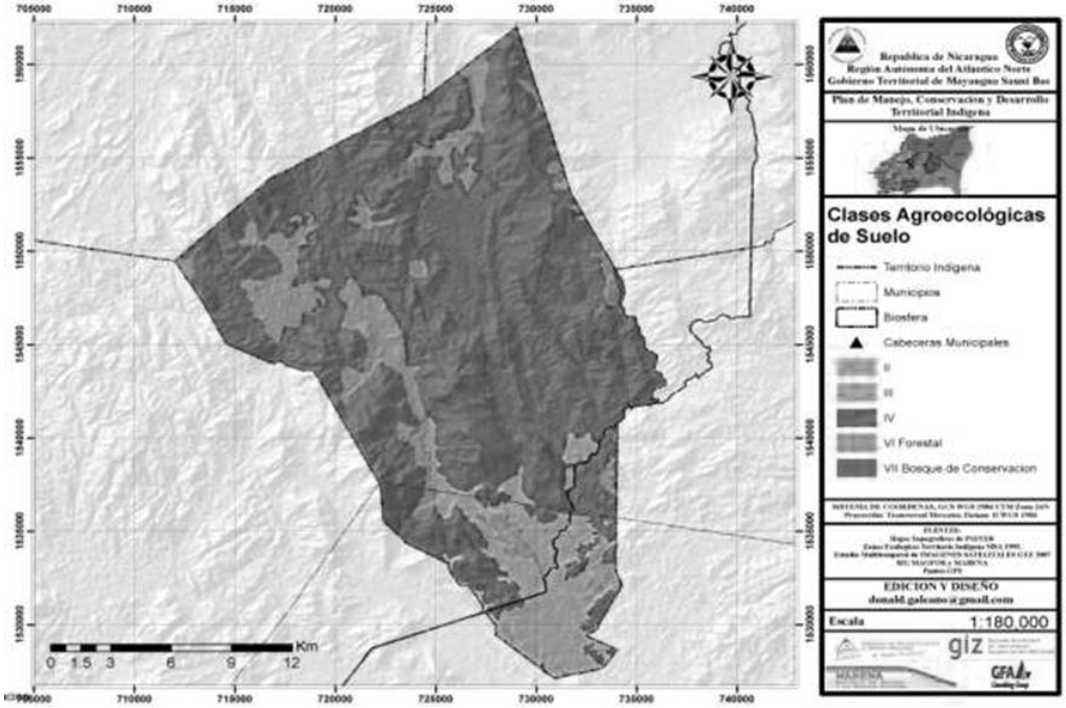


Image shows a map of soil types within Miskitu lands from Plan de Manejo, Conservación, y Desarrollo Territorial Indígena

The Miskitu is a large nation with several areas run by tribal territorial governments in Nicaragua. The Territorial Government of Mayangna Sauni Bas together with the Nicaraguan Ministry of Environment and Natural Resources and the German Technical Cooperation published "The Administration, Conservation, and Indigenous Territorial Development Plan" in 2011. [9]

This grey literature source is especially valuable as it comes from indigenous leadership and speaks directly to the people's concerns for their land today and in the future. The multi-faceted resource describes the evolution of the Plan, lists other indigenous territories in the region, analyzes biological diversity, and outlines local tribal governance structures. Maps included in the Plan range from hydrology to annual temperatures to changes in land usage since 1980.

The establishment of territorial tribal governments flow from the Inter-American Court of Human Rights' 2001 decision regarding *Mayagna (Sumo) Community of Awas Tingni v. Nicaragua*. As described by the University of Arizona College of Law, it "is the first case in which an international tribunal with legally binding authority has found a government in violation of the collective land rights of an indigenous group, setting an important precedent for the rights of indigenous peoples in international law." [10] The court's decision obligated the state of Nicaragua to secure land titles not only for the Awas Tingni community but all indigenous peoples within the country. This decision, together with Nicaragua's Law 25 about regional autonomy[11] and Law 445 about communal land ownership,[12] are changing the map and governmental structures, which the Portal will track via resources like this one.

Impact

In 2015, the Portal was honored with an ALA (American Library Association) Choice Reviews 2015 Outstanding Academic Title (Internet resource). Hard feedback in the way of web metrics was gathered by the Library of Congress to examine the usage of the Indigenous Law Portal in the two years since it launched. The Portal represents a service for and about diverse communities and the table below confirms that users come from all continents of the globe with the great majority coming from the United States. The following world map illustrates the concentration of users by country:

More specifically, the top 10 country users are on four continents. Users hail from 174 nations; as a point of reference, the United Nations has 193 member states.[13]

Top 10 Countries	# of users	% of users
United States	53,027	87.10%
Canada	1,292	2.10%
United Kingdom	840	1.40%
India	497	0.80%
Australia	493	0.80%
Germany	486	0.80%
Japan	353	0.60%
Poland	259	0.40%
France	216	0.40%
China	194	0.30%

What these users do and seek is illuminated below. The average user looks at two pages of the Portal during a visit. Users have downloaded close to 600 different documents, almost entirely grey literature products. The three most downloaded documents are classification schedules yet these are less than 17% of the 30 most popular downloads. The majority of downloads are constitutions and bylaws from a variety of tribes. Such feedback is encouraging.

Indigenous Law Portal Usage Statistics	
June 2014-December 2016	
Total Page Views	126,045
Total Visits	60,849
Total Hours	1842
Total File Downloads	3,340

Conclusion

The story of the Indigenous Law Portal is in good measure the story of diverse grey literature sources being leveraged via the classification, which is itself a grey literature resource. The librarians work collaboratively to build the Portal but it takes more than academic training and good will to create a space for indigenous law among humanity's legal systems. With respect to the Portal, the author and her colleagues at LC look to other professionals and especially indigenous peoples to voice their thoughts. All contributions can be shared with the world via the Portal in a virtuous circle of legal knowledge.

References

Amelung, Richard. Interview by author. Washington, DC, August 10, 2016.

Brammer, Robert. Interview by author. Washington, DC, July 28, 2016.

Cristan, Ana. Interview by author. Washington, DC, August 1, 2016.

Dechman, Elizabeth. Interview by author. Washington, DC, August 1, 2016.

Gheen, Tina. Interview by author. Washington, DC, August 1, 2016.

Goldberg, Jolande. Interview by author. Washington, DC, August 3, 2016.

Gonzalez, Jennifer. Interview by author. Washington, DC, August 4, 2016.

James, Jill. Interview by author. Washington, DC, August 3, 2016.

Lapidus, Noah. Interview by author. Washington, DC, August 3, 2016.

Neil, David. Gonzalez, Jennifer. Interview by author. Washington, DC, August 1, 2016.

Nolan, Jacqueline. Interview by author. Washington, DC, August 11, 2016.

Shaffer, Roberta. Interview by author. Washington, DC, August 14, 2016.

ENDNOTES

[1] Martin Nakata, Alex Byrne, Vicky Nakata, and Gabrielle Gardiner, "Indigenous Knowledge, the Library and Information Service Sector, and Protocols" in *Australian Indigenous Knowledge and Libraries,* ed. Martin Nakata and Marcia Langton (Sydney: UTSePress, 2005), 16. https://opus.lib.uts.edu.au/bitstream/10453/19531/1/Libraries%20and%20Indigenous%20Knowledge.pdf

[2] Linc Kesler, "Aboriginal Identity & Terminology." Indigenous Foundations. Accessed September 15, 2016. http://indigenousfoundations.arts.ubc.ca/home/identity/aboriginal-identity-terminology.html

[3] Alejandra Gámez Espinosa, *Pueblos Indígenas del México Contemporáneo: Popolocas,* 2006, https://www.gob.mx/cms/uploads/attachment/file/12580/popolocas.pdf

[4] Instituto Nacional de Lenguas Indígenas, Presentación de la Ley General de Derechos Lingüísticos de los Pueblos Indígenas, 2015, http://www.inali.gob.mx/es/difusion/presentacion-de-la-ley-general-de-derechos-linguisticos-de-los-pueblos-indigenas.html

[5] Stacey Caron Castillo Lewis, "Incidencia de la migración en la pérdida de la identidad cultural garífuna," thesis from the Universidad de San Carlos de Guatemala, 2005, https://perma.cc/FTT3-HP8P

[6] UNICEF, Agencia Española de Cooperación Internacional para el Desarrollo AECID, Fundación para la Educación en Contextos de Multilingüismo & Pluriculturalidad FUNPROEIB Andes, *Atlas sociolingüístico de Pueblos Indígenas de América Latina Fichas nacionales tomo II*, 2009, https://www.unicef.org/honduras/tomo_2_atlas.pdf

[7] Movimiento Indígena Lenca de Honduras MILH and Foro Internacional de Mujeres Indígenas FIMI, *Diagnóstico participativo en el pueblo indígena Pech Honduras C.A.*, no year, https://perma.cc/9RGT-DVDJ

[8] Ibid. No page number. Quotation as published ""Los líderes caciques dicen: "hay que regresar aplicar castigos severos como aplicaban nuestros antepasados, para que se den cuenta los hombres de hoy lo duro que son los castigos y no maltraten a sus compañeras de hogar." A continuación nos describieron un castigo que lo aplicaban que consta de hacer una fogata y poner en la parte alta palos en forma de cortina (Tabanco) y subirlo ahí para que trague todo el humo y si no es capaz de aguantar todo el tiempo necesario pues le dan una paliza hasta dejarlo inmóvil. "Este es un castigo que lo aplicaban nuestros antepasados" explica el cacique."

[9] Gobierno Territorial Mayangna Sauni Bas, Ministerio del Ambiente y Recursos Naturales, Cooperación Técnica Alemana (GTZ), *Plan de Manejo, Conservación, y Desarrollo Territorial Indígena*, 2011, https://perma.cc/3HUD-6VNS

[10] University of Arizona James E. Rogers College of Law, "IPLP Outreach Awas Tingni Case," http://law2.arizona.edu/Depts/iplp/outreach/awas.cfm

[11] Asamblea Nacional de la República de Nicaragua, Reglamento a la Ley No. 28 Estatuto de Autonomía de las Regiones de la Costa Atlántica de Nicaragua, 2003, https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/92444/107676/F-622526820/NIC92444.pdf

[12] Asamblea Nacional de la República de Nicaragua, Ley No. 445 Ley del Régimen de Propiedad Comunal de los Pueblos Indígenas y Comunidades Étnicas de las Regiones Autónomas de la Costa Atlántica de Nicaragua y de los Ríos Bocay, Coco, Indio y Maíz, 2003, http://www.poderjudicial.gob.ni/pjupload/costacaribe/pdf/Ley_445.pdf

[13] United Nations. "Overview." Accessed September 6, 2016. http://www.un.org/en/sections/about-un/overview/index.html

Transition to Open Access and its Implications on Grey Literature Resources

Snježana Ćirković, Austrian Academic Library Consortium, Austria

Abstract
This paper describes the transition to Open Access and its implications on grey literature resources. In this paper we will present current Open Access models, known as "offsetting deals", which main intention is to avoid "double dipping". This part will also review the role of library consortia in this process, as well as current Open Access policies in Europe.
The second part of this paper will explore the role of grey literature in transition process to Open Access. Grey literature is an important source of original research and up to date information, although the lack of peer review and formal publication standards must be taken into account during an evaluation process. Grey literature plays an important role in the rapid and timely distribution of in-depth, recent, scientific and technical information, and also provides access to a broad range of information and often contains new ideas. Research that is not published in journals but available in other formats (such as reports, theses or conference proceedings) is often more detailed, more recent and sometimes more rapidly disseminated. Due to the competitive and time consuming nature of publishing in peer-reviewed academic journals, some research may never make into journals and would, therefore, be inaccessible to interested parties without the grey literature. We will present possible ways of increasing the visibility of grey literature repositories, their inclusion in open access databases, and how to connect these institutional grey literature repositories with current research information systems.

Introduction
The term Open Access (OA) broadly refers to the accessibility of various documents, research outputs and other literature free of any charges and having fewer restrictions on their use, citation, reproduction and onward transmission.

In recent years Open Access has become a new focus of information resource innovation. Whereas premium scholarly content was formerly available only through expensive print subscriptions or proprietary databases, the open access movement promises to realize one of the fundamental aspirations of the public library movement, which is to make information in all its forms available to any interested citizen, without regard for ability to pay.

In its most elemental form, Open Access can be defined as providing free access for all Internet users to materials that have traditionally been published in scholarly journals. The more formal phrasing from the Scholarly Publishing and Academic Resources Coalition (SPARC), modeled after the Budapest Open Access Initiative's definition, states:

"By Open Access, we mean the immediate, free availability on the public internet… permitting any user to read, download, copy, distribute, print, search or link to the full text of these articles, crawl them for indexing, pass them as data to software or use them for any other lawful purpose."[1]

In line with the Berlin Declaration on Open Access of 22 October 2003, Science Europe defines Open Access as: "unrestricted, online access to scholarly research publications (including books, monographs and non-traditional research materials) for reading and productive re-use, not impeded by any financial, organisational, legal or technical barriers".[2]

[1] http://www.budapestopenaccessinitiative.org/read
[2] http://www.scienceeurope.org/policy/working-groups/open-access-to-scientific-publications/

Fig. 1: Benefits of Open Access (Source: SPARC, right to research, Kingsley & Brown 2012)

1. Open Access Policies

The growth of OA has been dramatic, but nowhere has this been more obvious than in the area of OA policies. This can only be understood when OA is seen in the context of wider political and policy issues. In this paper we will give an overview of some most significant European OA policies and policy statements starting with the Budapest Open Access Initiative.

1.1. The Budapest Open Access Initiative

In February 2002 the Budapest Open Access Initiative outlined two complementary routes to OA. The 'green' route was for researchers to deposit their peer-reviewed papers in open, interoperable electronic archives or repositories The 'gold' route was to publish in OA journals. In either case, there should be no financial, legal, or technical barriers between the author and the interested reader.

Definition of OA according to the Budapest Open Access Initiative - By "open access" to this literature, we mean its free availability on the public Internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited."[3]

The 16 original signatories of the Budapest Open Access Initiative . As of 20th January 2017, more than 6014 individuals and 916 organizations have signed it.

1.2. The Bethesda Statement on Open Access Publishing

Along with the *Budapest Open Access Initiative (BOAI)* and *the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*, the *Bethesda Statement* established "open access" as the term to describe initiatives to make information more widely and easily

[3] Budapest Open Access Initiative, http://www.budapestopenaccessinitiative.org/read

available. On 11 April 2003, the Howard Hughes Medical Institute held a meeting attemded by 24 people to discuss better access to scholarly literature. The group created a definition of an open access journal as one which grants a "free, irrevocable, worldwide, perpetual right of access to, and a license to copy, use, distribute, transmit, and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship" and from which every article is "deposited immediately upon initial publication in at least one online repository".[4]

1.3. The Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities in October 2003

The **Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities** is a major international statement on open access and access to knowledge. Following the Budapest Open Access Initiative in 2002 and the Bethesda Statement on Open Access Publishing in 2003, the Berlin Declaration was a third influential event in the establishment of the open access movement. Peter Suber has referred to the three events combined as the "BBB definition" of open access (Suber, 2012).

The declaration was drafted at October 2003 conference held by the Max Planck Society and the European Cultural Heritage Online (ECHO) project. More than 120 cultural and political organizations from around the world attended. The statement itself was published on 22 October 2003. Acknowledging the increasing importance of the Internet and the previous discussions on the need for open access, it offered the following definition of an open access contribution:

"Open access contributions must satisfy two conditions: The author(s) and right holder(s) of such contributions grant(s) to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they do now), as well as the right to make small numbers of printed copies for their personal use. A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in an appropriate standard electronic format is deposited (and thus published) in at least one online repository using suitable technical standards (such as the Open Archive definitions) that is supported and maintained by an academic institution, scholarly society, government agency, or other well-established organization that seeks to enable open access, unrestricted distribution, inter operability, and long-term archiving."[5]

It also encouraged researchers and institutions to publish their work in accordance with these principles, advocate for open access and help in the development and assessment of open access related tools and measures.

1.4 Policy Guidelines for the Development and Promotion of Open Access by Alma Swan, commissioned by UNESCO, 2012

These Guidelines provide an account of the development of Open Access, why it is important and desirable, how to attain it, and the design of effective policies. Open Access is a new way of disseminating research information, made possible because of the World Wide Web.
The benefits of Open Access are summarized as follows:

- Open Access improves the speed, efficiency and efficacy of research
- Open Access is an enabling factor in interdisciplinary research
- Open Access enables computation upon the research literature
- Open Access increases the visibility, usage and impact of research
- Open Access allows the professional, practitioner and business communities, and the interested public, to benefit from research

[4] http://legacy.earlham.edu/~peters/fos/bethesda.htm
[5] https://openaccess.mpg.de/Berlin-Declaration

- As Open Access has grown, new business models have been developed – for journal publishing, for Open Access repositories, book publishing and services built to provide for new needs, processes and systems associated with the new methods of dissemination [6]

1.5 Science Europe Principles on Open Access to Research Publications, April, 2013

Principles on the Transition to Open Access to Research Publications adopted in April 2013 by the Science Europe Member Organisations share the following view:

- Publication and dissemination of results are an integral part of the research process. The allocation of resources within the research system must take this into account.
- Open Access to the published results of publicly-funded research will have huge value for the research community and will offer significant social and economic benefits to potential users in industry, charitable and public sectors, to individual professionals, and to the general public.
- Open Access, as defined in the Berlin Declaration, is not only about the right of access, but also about the opportunity to re-use information with as few restrictions as possible, subject to proper attribution.
- The common goal of Science Europe Members is to shift to a research publication system in which free access to research publications is guaranteed, and which avoids undue publication barriers. This involves a move towards Open Access, replacing the present subscription system with other publication models whilst redirecting and reorganising the current resources according to Grey Literature.
- Science Europe is committed to playing a role in accomplishing the transition to Open Access as quickly as possible, in an efficient and sustainable way, and thus avoiding unnecessary costs. This transition process must be as co-ordinated and transparent as possible."[7]

1.6 ROARMAP (The Registry of Open Access Repository Mandates and Policies)

The ROAR is a searchable international Registry of Open Access Repositories indexing the creation, location and growth of open access institutional repositories and their contents. It was created by ePrints at University of Southampton in 2003. To date, over 3,000 institutional and cross-institutional repositories have been registered in the ROAR.

The ROAR's companion database, the Registry of Open Access Repositories Mandatory Archiving Policies (ROARMAP), is a searchable international registry charting the growth of open access mandates adopted by universities, research institutions and research funders that require their researchers to provide open access to their peer-reviewed research article output by depositing it in an open access repository.

1.7 Amsterdam Call for Action on Open Science

The Amsterdam Call for Action on Open Science is a document that advocates for "full open access for all scientific publications", and endorses an environment where "data sharing and stewardship is the default approach for all publicly funded research". The Amsterdam Call for Action on Open Science was first produced as a draft at an Open Science meeting organized by the Dutch Presidency of the Council of the European Union on 4-5 April 2016, in Amsterdam.

The finalised Call for Action was input to the Competitiveness Council on the 27th of May, led by the Dutch State Secretary for Education, Culture and Science Sander Dekker. In the main results of this meeting a reference was made to the Call for Action in the document called "Outcome of the Council Meeting". It was also referred to in the Draft Council conclusions on the transition towards an Open Science system, point 3 under the section on Open Science. In the press comments on the meeting of the Competitiveness Council, the focus was mainly on the fact that European leaders call for 'immediate' open access to all scientific papers by 2020.[8]

[6] http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/policy-guidelines-for-the-development-and-promotion-of-open-access/

[7] http://www.scienceeurope.org/policy/working-groups/open-access-to-scientific-publications/

[8] (https://enGreyLiteratureish.eu2016.nl/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science)

2. Transition to Open Access – Global level view

The Max Planck Digital Library has put forward a study on the transformation of the subscription-driven system for scientific publications to an Open Access model. For the first time, quantitative parameters were presented showing that the liberation of scholarly literature is possible at no extra costs. "According to market analyses, annual turnovers of academic publishers amount to approximately EUR 7.6 billion. This money comes predominantly from publicly funded scientific libraries as they purchase subscriptions or licenses in order to provide access to scientific journals for their customers. Since more than a decade the Open Access movement, in which the Max Planck Society plays a major role, has been demanding free and immediate access to the results of academic research on the Internet.

Open Access publishers ensure their financial sustainability through charging publication fees: Not readers but rather authors or their institutions or funders are supposed to pay for publications. While numerous publishers have already adopted an Open Access business model during the last few years, the share of openly available scientific articles is still only at a level of some 13%." (Schimmer, 2015)

"We need to create an efficient and widely budget-neutral transition which offers fresh incentives to traditional publishers to cooperate and transfer established journals to Open Access," says Martin Stratmann, President of the Max Planck Society.

The study which has been presented by the Max Planck Digital Library is investigating the question whether the previously used subscription budgets would be sufficient to fund the Open Access publication charges and thus bring about a complete transition of academic publishing. The paper, entitled *Disrupting the subscription journal's business model for the necessary large-scale transformation to open access,* concludes that such a transition would be possible at no extra costs. "An internationally concerted shifting of subscription budgets is possible at no financial risk, maybe even at lower overall costs", says main author Ralf Schimmer.[9]

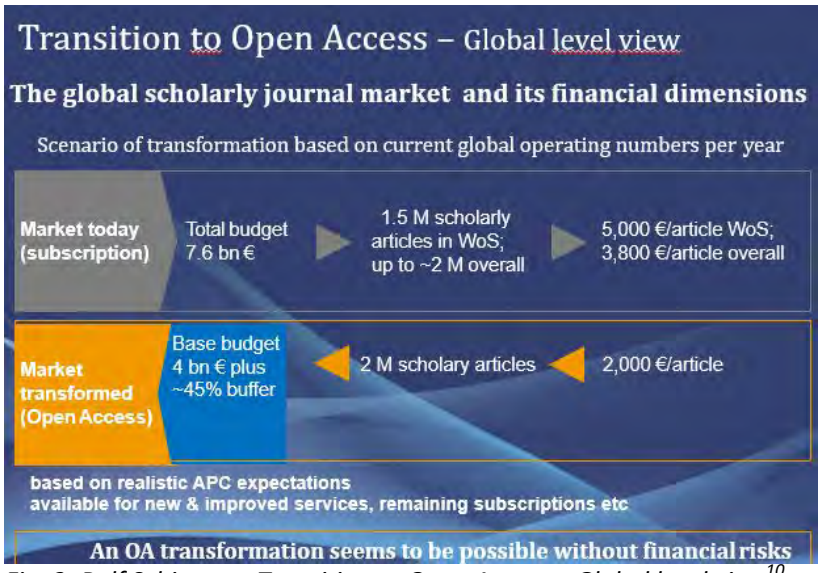


Fig. 2: Ralf Schimmer, Transition to Open Access – Global level view[10]

3. Transition to Open Access in Austria

The Austrian Science Fund (FWF) supports Open Access both conceptually and financially since 2003. The Council of Research and Technology Development (RFTE) explicitly affirmed the concept of Open Access in its 2009 *Strategy 2020*. The 50 member organisations of the Open Access Network Austria (OANA), initiated by Universities Austria (uniko) and the FWF and supported by the Federal Ministry of Science, Research and Economy (BMWFW), are advancing Open Access significantly. Besides, Open Access and Open Science are core

[9] http://pubman.mpdl.mpg.de/pubman/item/escidoc:2148961:7/component/escidoc:2149096/MPDL_OA-Transition_White_Paper.pdf

[10] http://liber2016.org/wp-content/uploads/2015/10/1400-1420_Schimmer_Open_Access_2020.pdf

elements of the Alliance for Responsible Science, Open Innovation, the Digital Austria platform, and the ERA Roadmap 2016 of the Austrian Federal Government. Based on these principles, efforts should be made to achieve the following goal:

"By 2025, a large part of all scholarly publication activity in Austria should be Open Access. In other words, the final versions of most scholarly publications (in particular all refereed journal articles and conference proceedings) resulting from the support of public resources must be freely accessible on the Internet without delay (Gold Open Access). This goal should be pursued by taking into account the different disciplinary practices and under consideration of the different disciplinary priorisations of Open Access. The resources required to meet this obligation shall be provided to the authors, or the cost of the publication venues shall be borne directly by the research organisations. The necessary funding must be brought in line with the overall funding priorities for research."[11]

In Austria, a number of closely collaborating bottom-up initiatives have been established through the FWF, the Open Access Network Austria (OANA), the Austrian Academic Library Consortium (KEMÖ), and e-infrastructures Austria. Given appropriate research-policy-based support, these initiatives could serve as a basis for positioning Austria as an innovation leader for Open Access in Europe and throughout the world.

3.1. Austrian Academic Library Consortium (KEMÖ) in a Nutshell

The Austrian Academic Library Consortium currently consists of 58 institutional libraries, among them 18 universities, 8 private universities, 18 universities of applied sciences, and 14 other institutions. The central purpose of the KEMO is:

- the coordinated acquisition and management of electronic resources (primarily databases and e-journals), and the usage rights of these resources within the consortia
- coordinated resource administration
- a unified voice to represent the consortia members both internationally and within library consortia, such as GASCO (German, Austrian and Swiss Consortia Organisation), ICOLC (International Coalition of Library Consortia), and other interest groups,
- joint collaboration for national and international Open Access initiatives. The Head Office of Kooperation E-Medien (Austrian Academic Library Consortium) is a department of the Austrian Library Network and Services Ltd (Österr. Bibliothekenverbund und Service GmbH), which also oversees operative support.[12]

3.2. FWF Austrian Science Fund

The Austrian Science Fund (FWF) is Austria's central funding organization for basic research. The purpose of the FWF is to support the ongoing development of Austrian science and basic research at a high international level. In this way, the FWF makes a significant contribution to cultural development, to the advancement of the knowledge-based society, and thus to the creation of value and wealth in Austria.

3.3. OANA – Open Access Network Austria

Open Access Network Austria (OANA) was established in 2012 as a joint activity under the organisational umbrella of the Austrian Science Fund (FWF) and Universities Austria (UNIKO). The network ties in with the uniko Open-Access Recommendation (12.01.2010) and the FWF-Position Paper (17.01.2012) and develops specific recommendations for the implementation of Open Access in Austria:

- Coordination of and recommendations for the Austrian OA-task/ activities of research institutions, funding organisations and research policies (including international developments)
- Positioning towards the information providers (mainly publishing houses)
- Contact persons and resource of information for scientists, research institutions and (research-) policies. [13]

[11] https://zenodo.org/record/34079#.WITSHlz3gfI
[12] https://konsortien.at/default-en.asp

[13] http://www.oana.at/

3.4. UNIKO - Universities Austria

Universities Austria is a non-profit association under private law. Its purpose is to assist the Austrian universities in the fulfilment of their tasks and responsibilities and thus to foster scholarship and research.[14]

3.5. Recommendations for the Transition to Open Access in Austria

In order to achieve the goal that by 2025, a large part of all scholarly publication activity in Austria should be Open Access, a set of 16 coordinated measures are recommended:

(1) Introduce Open Access policy
(2) Create cost transparency
(3) Reorganise publishing contracts
(4) Introduce publication funds
(5) Reorganise publication venues
(6) Merging the publication infrastructure
(7) Support international cooperation
(8) Provide start-up capital
(9) Registration of repositories
(10) Support self-archiving
(11) Offer training programmes
(12) Acknowledging Open Access / Open Science
(13) Expand the scope of the copyright reform of 2015
(14) Opening the inventories
(15) Monitoring during implementation
(16) Set targets for Open Science

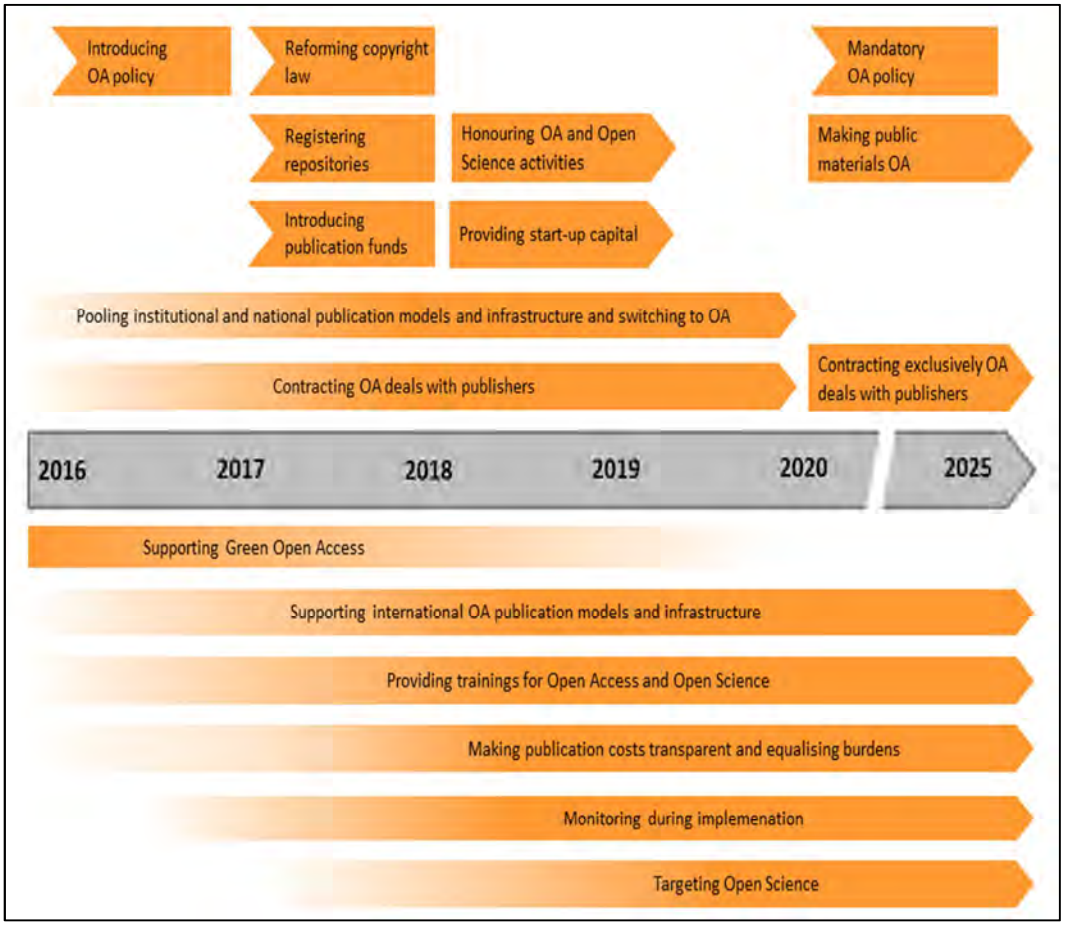


Fig. 3: Timeline of the Open Access Recommendations

[14] https://uniko.ac.at/index.php?lang=EN

3.6 Open Access Deals in Austria

Scholarly journals are the core of the current publication system. Scholarly journals are purchased by the libraries of research institutions in large packages, especially from large publishers, to enable their staff to access the publications. In Austria, libraries are currently purchasing packages exceeding 30 million Euros per year through the consortium network known as the Austrian Academic Library Consortium (KEMÖ). It is suggested that this model will be transformed to Open Access in three temporally coordinated steps:

Offsetting models - This term refers to the implementation of models that function on the following basis: payments for Open Access in subscription journals (Hybrid Open Access) should be offset against the subscription prices of the journals or should cause a high discount. KEMÖ and the FWF have concluded preliminary global agreements of this nature with publishers like IOP Publishing and Taylor & Francis in 2014. For 2017 KEMÖ closed three Offsetting deals, with IOP Publishing, Taylor&Francis and Emerald Publishing Group. Negotiations with other publishers are under way. However, these can only serve as entry-level models into a transitional phase.

Read & Publish models - In a mid-term step, attempts should be made to implement models that include access permission for the subscribing institutions as well as an Open Access publication option for scholars of the institutions. Such "Read & Publish models" were closed with Springer (Springer Compact 2016-2018) and with Royal Society of Chemistry (Read & Publish Deal 2017-2018).

Open Access service-based models - In the long term, i.e. from 2020 onward, contracts should be concluded in a manner that the price is no longer derived from the subscription package but from the costs of the individual published articles. Based on this suggestion the Max Planck Society assumes that the former resources would not only be sufficient for a model of this type, but would also permit competitive pricing which could reduce costs.[15]
Based on experience obtained in the Netherlands and in Austria, such substantial modifications of contracts, especially with large multinational enterprises, would be successful only if the Austrian Academic Library Consortium (KEMÖ) (a) was strengthened, and (b) negotiations were supported even more actively by the heads of research institutions.

4. The role of Grey Literature and inclusion the GL resources in Open Access Databases

Grey literature is commonly described as a field in library and Information science that deals with the production, distribution, and access to multiple document types produced on all levels of government, academics, business, and organizations in electronic and print formats not controlled by commercial publishing, i.e. where publishing is not the primary activity of the producing body. Examples of grey literature include patents, technical reports from government agencies or scientific research groups, working papers from research groups or committees, white papers, and preprints.

> "Grey literature stands for manifold document types produced on all levels of government, academics, business and industry in print and electronic formats that are protected by intellectual property rights, of sufficient quality to be collected and preserved by library holdings or institutional repositories, but not controlled by commercial publishers i.e., where publishing is not the primary activity of the producing body " (Schöpfel, 2011)

In the time of fast information exchange, grey literature has an important role and offers to scientists faster and free access to the knowledge. We should use the opportunity to underline that role, by including grey literature resources in open access databases. In that sense, we will increase the visibility of grey literature. We should put our efforts to improve metadata standards and protocols of grey literature.

[15] https://zenodo.org/record/34079#.WITSHlz3gfI

Grey literature should become part of packages and different databases, but with no additional costs to the end-users. The most important task is to include grey literature in open access databases, in order to increase recognition of grey literature for open access to research, open science and knowledge transfer. Including the grey literature resources in the following Open Access Databases will increase their usage and visibility:

- DOAJ – Directory of Open Access Journals[16]
- Sherpa Romeo Database [17]
- OAPEN Library [18]
- OAFindr 1science [19]

In Austria there is a database "STICHWORT Bibliothek" which defines the grey literature as following: "Grey literature are publications which are not published in the traditional way, e.g. research reports, brochures, typescripts, business reports, project concepts, and more." The "STICHWORT Bibliothek" contains a relatively large proportion of Grey Literature. These are mostly research documents, which are not preserved otherwise.

5. Conclusions

Although the Open Access movement is one of the major subjects in the academic publishing in the last 15 years, we are still in the period of transition to the pure Open Access world. The necessary switch from the subscription based models to the pure OA models didn't happen yet. According to ROARMAP, there are more than 270 OA Policies only in Europe, but this switch done globally, so that we could step into the new OA era. Many European countries are making their efforts develop national strategies and policies and different models of scholarly publishing in this transition period. Austria is one of them. According to the *Recommendations for the Transition to Open Access in Austria*, "by 2025, a large part of all scholarly publication activity in Austria should be Open Access". In order to achieve this goal, one of the measures is developing transformation models for scholarly journals, known as *Offsetting Deals, Read&Publish Deals,* and in the last phase of transition, *Open Access Service-Based Models*. Grey Literature is the part of scholarly research, and it should be included in these suggested transformation models.

Open Access promotes transparency and public insight into scientific outcomes, and including the Grey Literature resources in OA Databases will increase the visibility and usage of Grey Literature. If we say that Open Access is a prerequisite for Open Science, where all research data will be open and transparent, the Grey Literature should be already included in these transformation models, which will increase the visibility, usage and significance of Grey Literature resources in the scholarly publishing.

[16] https://doaj.org/
[17] http://www.sherpa.ac.uk/romeo/index.php
[18] http://www.oapen.org/home
[19] http://www.1science.com/oafindr.html

References

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, http://oa.mpg.de/openaccess-berlin/berlindeclaration.html (accessed 20 January 2017)

Bethesda Statement on Open Access Publishing, http://legacy.earlham.edu/~peters/fos/bethesda.htm (accessed 20 January 2017)

Budapest Open Access Initiative, http://www.soros.org/openaccess/read.shtml (accessed 20 January 2017)

Farace, D. (2014). The GreyGuide Repository and Web-Access Portal: GreyNet's Response to the Pisa Declaration. *Conference on Grey Literature and Repositories: proceedings 2014: the Value of Grey Literature in Repositories* [online]. Prague: National Library of Technology. Retrieved from: http://repozitar.techlib.cz/record/808/files/idr-808_4.pdf

Policy Guidelines for the Development and Promotion of Open Access by Alma Swan, commissioned by UNESCO, 2012, http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/policy-guidelines-for-the-development-and-promotion-of-open-access/ (accessed 20 January 2017)

Prosser, D. C. (2007). Public Policy and the Politics of Open Access. *LIBER Quarterly, 17*(2). doi:10.18352/lq.7877

Science Europe Principles on Open Access to Research Publications, April, 2013, http://www.scienceeurope.org/policy/working-groups/open-access-to-scientific-publications/ (accessed 20 January 2017)

Schimmer, R., Geschuhn, K. K., & Vogler, A. (2015*). Disrupting the subscription journals' business model for the necessary large-scale transformation to open access*. doi:10.17617/1.3. , http://hdl.handle.net/11858/00-001M-0000-0026-C274-7 (accessed 20 January 2017)

Schöpfel, J. (2010). Towards a Prague Definition of Grey Literature. *Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues. Prague, 6-7 December 2010, Dec 2010, Czech Republic.* pp.11-26.Retrieved from: http://hdl.handle.net/10068/700015

Suber, P. (2012). *Open access*. Cambridge, Mass: MIT Press.

Developing, linking, and providing access to supplemental genetics dataset vcf files

Plato L. Smith II, University of Florida (Libraries), United States
Lauren McIntyre, University of Florida (Microbiology), United States

Abstract
This conference proceeding paper is the written version component of the data panel discussion on developing a dataset collection using Zenodo for a professor in the Department of Molecular Genetics & Microbiology at the University of Florida.

An internal University of Florida George A. Smathers Libraries Strategic Opportunities Program (SOP) grant award provided support for the creation and development of an initial supplemental datasets digital collection of large, static variant call format (vcf) in zenodo. The "*Documenting a Genomics Variant Files Data Management: Developing Research Data management (RDM) workflows and providing research data access via HPC"* project inspired this paper. The large vcf datasets used for this project ranged from 34 megabytes to 43 gigabytes. The researcher needed to (1) develop a data repository for supplemental datasets vcf files too large for attachment as supplemental data files for journal submissions, (2) provide digital object identifiers (DOIs) for all vcf dataset files, and (3) link the supplemental vcf dataset files to the journal article via the vcf doi. These three outcomes were accomplished during phase 1 (June 2016 – December 2016) of this project and presented at the GL18. Phase 2 (January 2017 – June 2017) of this project includes performing (1) a dataset reproducibility interview, (2) an open archival initiative protocol for metadata harvest (OAI-PMH) from Zenodo to the University of Florida institutional repository (IR@UF), and (3) developing a similar use case project for researchers in UF/IFAS Nature Coast Biological Station (NCBS).

Introduction
Data continuously extended, stored, and consulted is useful to science (*As We May Think*, Bush, 1945). Data is useful to science if data is accessible, discoverable, and reproducible. This project allows researchers to access, discover, share, and cite large variant call format (VCF) datasets from a data repository. This project can be as a use case scenario for articulating, demonstrating, and detailing the data lifecycle.

Data generation, raw data files are often text files of large sizes. These raw data- effectively what comes off the machine have been a focus for many in terms of collection and preservation. While there has been a huge effort expended in this area for gene expression data, there has been much less effort for capturing and storing DNA based variant data, and mass spectrometry based metabolomics and proteomic data. As our capacity for data generation continues to expand, thoughtful discussion on what to keep, for how long, and where is an ongoing important public dialogue.

The gene expression and data collection as an example, array images (GEO), or sequence fastq files in the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) are the raw data. The linking of signal intensity, or number of reads to an identifier is the basis then for all subsequent analysis in the Gene Expression Omnibus (GEO). "Geo is a public functional genomics data repository supporting MIAME-compliant data submissions" (NCBI Resources, GEO). Updates, deposition of quantified data, and annotation of the features for the snapshot in time are possible. The quantified processed file is the basis for further analysis and is an intermediate file. Public access and long-term storage of primary data in GEO are important to the research community. The full list of features and the result of whatever analysis done by an individual scientist is then a results file. The results file is often relatively small and it can be included with publication. While not consistently encouraged among all scientific journals, the GSA journals do try to capture this final file and include as supplementary material. Individual scientists should think more carefully about how the inclusion of a full supplement helps their own research long term.

Complete data capture, access, storage and reuse enable current and future researchers to start/build research. It obviates the necessity of identifying exactly which file in the folder was the full and final result file. Multiple versions, trainee turnover, and manuscript submission contribute to subtle differences among files on the main analysis system. The final supplement is an opportunity to capture the full information about the analytical results of a particular experiment. This is particularly beneficial if there is any break in the work on any one project, where trainees may not directly overlap. There are, of course, more altruistic reasons for including such files with a paper. However, the main beneficiaries of such preservation are often the lab that generated the results and close collaborators of the project.

As well developed, though imperfectly executed, this data preservation, processing, and sharing paradigm is for gene expression data, for other omics data these steps are significantly more challenging. In attempting to solve these challenges, we present one promising option for scientists to consider, partnerships with their institutional libraries. We describe one scenario here and invite the community to dialogue with us about these ideas.

For some population studies such as the 1000 genomes, special repositories are searchable have been created (). There is no consensus or centralized repository for all pieces and variant information. The affordability of sequencing efforts increases other populations sequencing. SRA (Sequence Read Archive) allows deposits of raw data in the form of sequence reads. Theoretically, full genomes can be stored at NCBI. Sequencing information, data quality, and the deposit workflow of genomes in NCBI needs enhancement. One of the most useful pieces of intermediate data is the list of variants, usually in a standard .vcf format. Depending on the population size, the .vcf files can be quite large. They are certainly cumbersome for inclusion in a journal supplement. It is not at all clear that these are good targets for either NCBI or journal supplements- for one compelling reason- the utility of these files is time limited. After the passage of some time, perhaps ten years, our algorithms will be better, we would likely prefer to reanalyse the raw data. The argument for semi-permanent storage of raw data is compelling, for intermediate files, it is less so. In some cases, the computational time to regenerate files is negligible. In the case of a vcf file, this is less true as creating these files often takes weeks if not months of computational effort, as well as considerable human effort along the way. There is therefore a good reason to keep these files in the short to medium term. How to keep them and then make them accessible for both the initial research group and the larger public has some current possibilities.

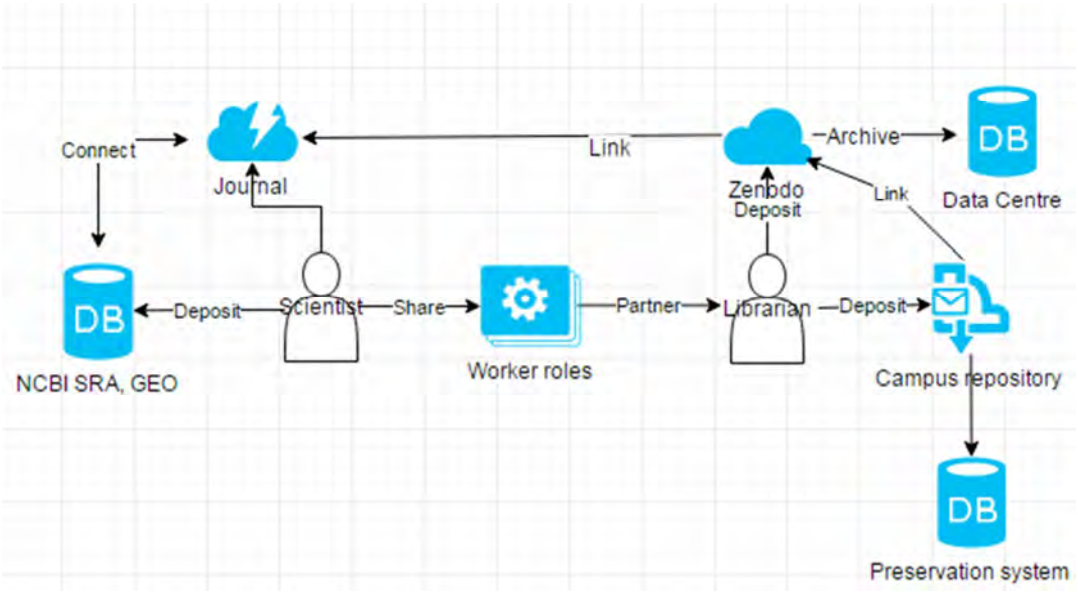


Figure 1. Raw data to SRA, results to journal, vcf to zenodo and IR@UF projected workflows

Researcher need for a vcf data repository

"We can generate terabytes of data about the genetics of population from a myriad of species. As a community, we have even made progress in developing unified data formats for reporting on observed genetic variation (vcf files). Yet, there is no national repository for this information. The libraries represent a transparent, public venue for sharing information on variants.

As a test project, the UFL library has collaborated with investigator Lauren McIntyre on a project in Drosophila (fruit fly). A population study of D. simulans funded by the NIH where ~200 different simulans genotypes were sequenced with 10x-20x coverage were analysed for variants. The resulting variant files have been deposited in the UFL library and links are included in the manuscripts currently under review." - UF Molecular Genetics & Microbiology Scientist

Using Zenodo as a vcf data repository

Zenodo is a public general research data repository that supports multiple forms of data from publication, poster, presentation, dataset, to image, video/audio, software, and lesson deposits up to 50 Gigabytes per record. The "Large 'static' processed data files" collection in zenodo enables the aggregation, representation, dissemination, and preservation of large, vcf datasets that exceeded the file size limits for supplemental data files for publications. These vcf datasets have digital object identifiers (DOI); and BibTex, CSL, DataCite, Dublin Core, JSON, MARCXML and Mendeley export features. The vcf datasets collection are harvestable via Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) using the Zenodo Harvesting API for vcf datasets collection. The datasets in zenodo are guaranteed access and preservation for at least 20 years. Zenodo support provided an excellent response to request for information on the access, storage, and preservation of the vcf datasets stored in zenodo. The following preservation information about datasets stored in zenodo is useful to the GL18 and DataONE Users Group communities.

'Zenodo stores its data as part of CERN's disk-storage service EOS (see http://information-technology.web.cern.ch/services/eos-service). This same service is being used for High Energy Physics data storage that is being obtained from CERN's LHC. The data is stored in CERN Data Centre (see general overview of CERN's Data Center here: http://information-technology.web.cern.ch/about/computer-centre as well as and more detailed report here on bit-preservation practices: https://cds.cern.ch/record/2195937/files/iPRES2016-CERN_July3.pdf). Additionally on the legal note of data hosting: CERN premises (including the Data Centre) are located on an intergovernmental territory, which is exempt from the host country's local jurisdiction (the host countries for the data centre is Switzerland and France, but as said their legislation does not apply since CERN has a status similar to United Nations).

As for Zenodo's internal workflow for data preservation - we are partially implementing the OAIS model for data archiving (ISO 14721), and are working towards being fully compliant with the model later this year or early next year. In parallel to that we will be working in the near future on fulfilling the requirements to be compliant with the Data Seal of Approval. All records in Zenodo (in all Zenodo collections) are treated equally and are undergoing the same procedures and workflows.

We compute MD5 checksums for all data that is uploaded - this information is visible to the user after upload, as well as returned in the response from the REST API for verification by the user. The checksum comes from the aforementioned CERN's EOS service, which is also used by EOS (alongside other measures) to prevent "bit-rot". For the time being we only generate and extract metadata from software repositories published through our GitHub integration. The metadata is normalized in a sense that its structure and data types are conforming to a pre-defined JSON schema, which is consistent across all records.' – Zenodo Technical Support, 2016

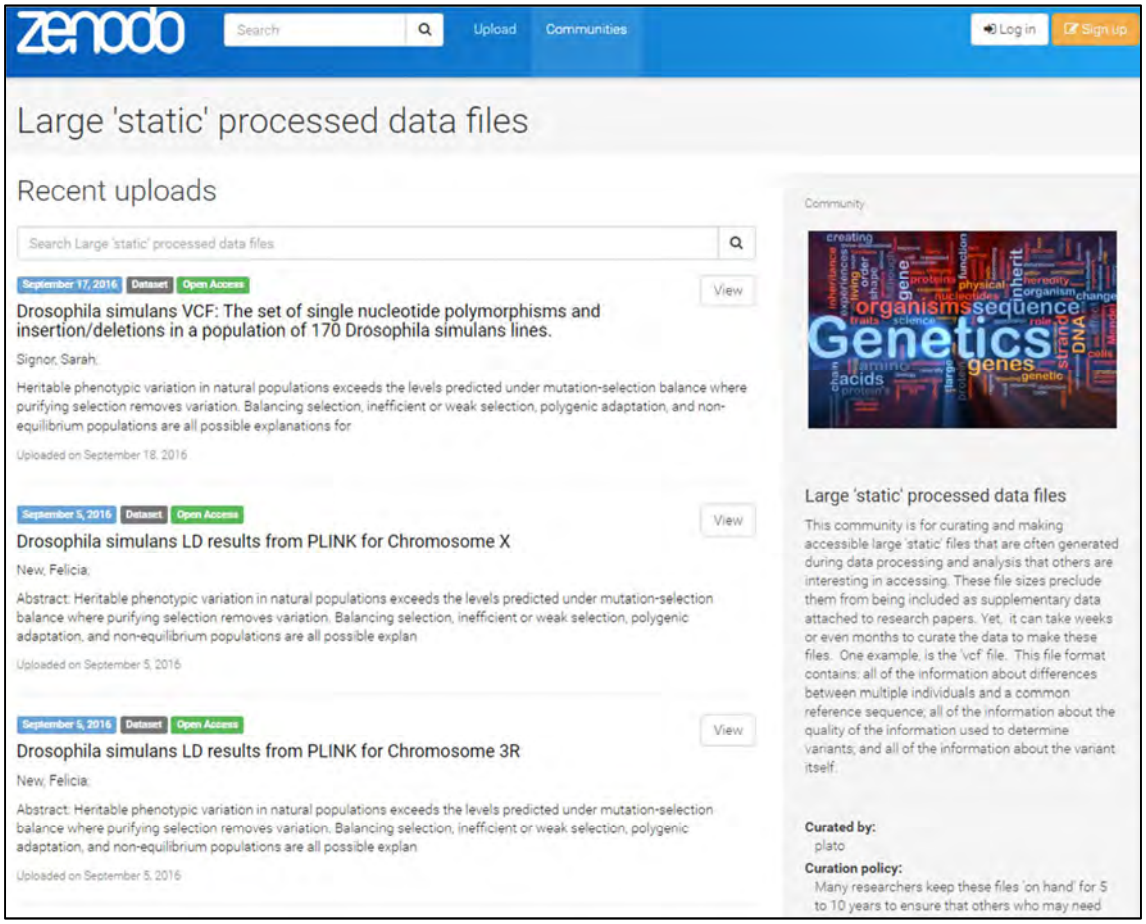


Figure 2. Large 'static' processed data files (vcf) repository for access, linking, and sharing

Benefits, features, and outcomes from this project

Some benefits and features from this project include but not limited to the following:

- Genetics project website in Open Science Framework (OSF) – Phase 1
- Genetic vcf datasets collection in zenodo - (Phase 1)
- DOI for vcf dataset files by DataCite
- Export vcf datasets in multiple formats (e.g. BibTeX, CSL (Citation Style Language) JSON Export, DataCite, Dublin Core, MARCXML, Mendeley)
- vcf datasets data collection harvestable via OAI-PMH API
- MD5 checksum performed on all uploaded data
- Zenodo partially implementing the OAIS model for data archiving (ISO 14721:2012) – Space data and information transfer systems - OAIS Reference Model – full compliance later this year or early next year
- Zenodo working in near future on fulfilling the requirements to be complaint with the Data Seal of Approval (DSA)
- Data access/preservation guaranteed for at least 20 years - FAQ
- Zenodo is US Department of Transportation (DOT) Public Access Plan conformant http://ntl.bts.gov/publicaccess/repositories.html

Acknowledgements
The authors acknowledge Allison Meryl Mores, Felicia New, Patrick Reakes, Ruth Isaacson, Genetics Editorial Office, and GreyNet/GL18.

References

[report] Consultative Committee for Space Data Systems. (2012). *Reference model for an Open Archival Information System (OAIS)* (Magenta Book CCSDS 650.0-B-1). Retrieved from http://public.ccsds.org/publications/archive/650x0m2.pdf

[website] IGSR: The International Genome Sample Resource. (2016). Retrieved from http://www.internationalgenome.org/wiki/Analysis/vcf4.0/

[journal article] Kurmangaliyev, Y. Z., Favorov, A. V., Osman, N. M., Lehmann, K., Campo, D., Salomon, M. P., Tower, J., Gelfand, M. S., & Nuzhdin, S. V. (2015). Natural variation of gene models in Drosophilia melanogaster. *BMC Genomics.* doi: 10.1186/s12864-015-1415-6

[report] National Institutes of Health. (2015). *National Institutes of Health Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research*. Retrieved from https://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf

[report] National Science Foundation Public Access Plan. (2015). *Today's Data, Tomorrow's Discoveries: Increasing Access to the Results of Research Funded by the National Science Foundation.* Retrieved from https://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf

[website] NCBI. (2016). SRA, Sequence Read Archive. Retrieved from https://www.ncbi.nlm.nih.gov/sra

National Repository of Grey Literature (NRGL)

NU¶
· SL

NRGL is

digital repository for grey literature

Free

online access

Features

Provider:
National Library of Technology
Prague, Czech Republic

Records:
over 400,000 records

Collection provenance:
Czech Republic

Partners:
over 130 organizations (Academy of Science, Public Research Institutions, Universities, State Offices, Libraries, NGOs etc.)

International Cooperation:
OpenGrey, OpenAire, ROAR, OpenDOAR, BASE

Goals

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

What else?

Conference on Grey Literature and Repositories
http://nrgl.techlib.cz/conference/

Informative Web pages
http://nrgl.techlib.cz

www.nusl.cz

NTK
Národní technická knihovna
National Library of Technology

NU¶
· SL
národní úložiště šedé literatury

nusl-letak-A4_en_01_cmyk.indd 1 10/10/2016 3:39:23 PM

Policy Development for Grey Literature Resources: An Assessment of the Pisa Declaration

Dobrica Savić, NIS-IAEA, Austria;
Dominic Farace and Jerry Frantzen, GreyNet International, Netherlands;
Stefania Biagioni and Carlo Carlesi, ISTI-CNR, Italy;
Herbert Gruttemeier and Christiane Stock, Inist-CNRS, France

Abstract
In the spring of 2014, a workshop took place at the Italian National Council of Research in Pisa[1]. The topic of this event dealt with policy development for grey literature resources. Some seventy participants from nine countries took an active part in the workshop – the outcome of which produced what is today known as the Pisa Declaration[2]. This fifteen point document arising from the input of those who attended the workshop sought to provide a roadmap that would help to serve diverse communities involved in research, publication and the management of grey literature both in electronic and print formats.
The Pisa Declaration has been translated and published in some twenty languages. They are all accessible online via the GreyGuide Repository[3] and Portal[4]. Currently, 140 information professionals from renowned organizations worldwide have endorsed this document[5].
In an effort to assess the impact that the Pisa Declaration has had during the last two years on the policy development for grey literature resources, an online survey among those who endorsed the document was carried out and their responses were analysed. Descriptive statistics and short summaries are used to describe the basic features of the data collected. They are combined with simple graphics that offer easier visual representation of the results achieved. Specific results of the survey analysis indicate those points in the Pisa Declaration that in varying degrees are of relevance and importance to grey literature, as well as points that need further attention and work. Although integral part of library and information management practice grey literature has its own peculiarities and needs that require special attention in order to reach its deserved level of importance in today's research and other activities.

Introduction
Since its publication in 2014, the Pisa Declaration on Policy Development for Grey Literature Resources has been endorsed by 140 signatories from 74 organizations in 30 countries worldwide. This Declaration has since been translated from the original English text into 20 languages and has come to be termed as the 'roadmap for grey literature in the 21st Century'. Now two years on, it is opportune to assess the impact this document has had on library and information practice. It is to this end that an online survey was conducted among its signatories, the results of which are found here recorded.

Survey sample size and population
Online questionnaire-based Pisa Declaration survey was created and placed on the SurveyMonkey on 25 April 2016. It consisted of 10 multiple choice questions with some of them offering a possibility to leave additional comments. The survey was designed in English language only. First survey replies were received on 30 May and the survey was closed on 18 July 2016. Figure 1 shows the survey response volume and the time distribution.
Requests for completion of the survey were sent to all 133 Pisa Declaration signatories, out of which 60 responded. This marked a 45% response rate.
Generally speaking there are two types of surveys - Surveys distributed internally, such as this one, since it was distributed only to a pre-set group of individuals, and external surveys distributed to wider audience, such as potential customers or general public.

[1] http://eventi.isti.cnr.it/index.php/en/programme-grey
[2] http://www.greynet.org/images/Pisa_Declaration,_May_2014.pdf
[3] http://goo.gl/72yexP
[4] http://greyguide.isti.cnr.it/
[5] http://greyguiderep.isti.cnr.it/pisadecla/listaiscritti.php?order=name

According to SurveyGizmo.com, internal surveys will generally receive a 30-40% response rate on average, compared to an average 10-15% response rate for external surveys.[6] Achieved response rate of 45% with the Pisa Declaration survey is therefore, regarded as above the average.

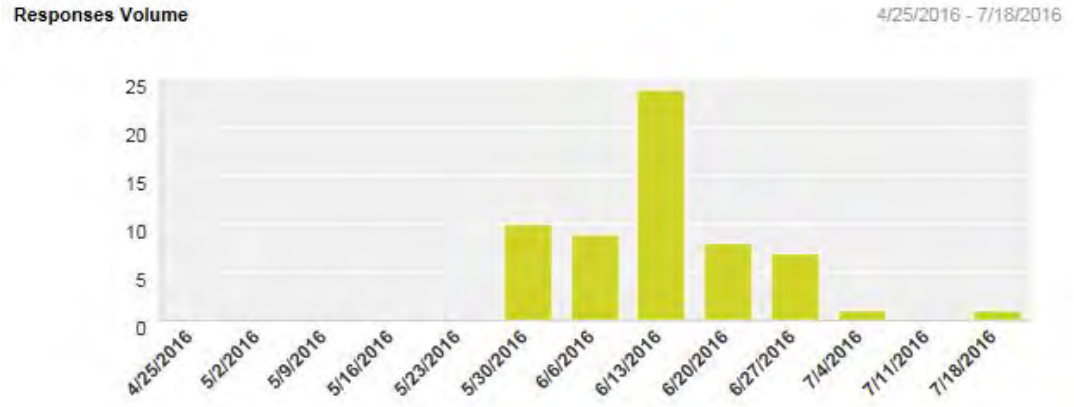


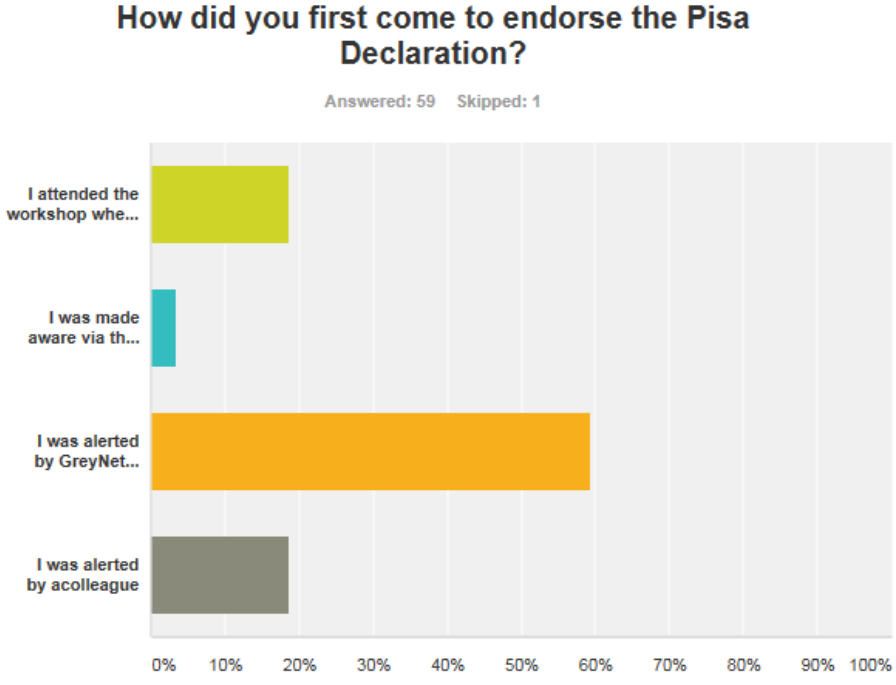
Figure 1: Response Volume

Question 1: How did you first come to endorse the Pisa Declaration?

The goal of this question was to find out the way respondent found out about the Pisa Declaration. Out of 60 respondents, 59 answered this question. One respondent skipped this question and left an interesting comment that he/she does not remember how it came out to endorse the Pisa Declaration.

Replies to this question offer some interesting conclusions. First of all, the website is not the best channel for promoting, grabbing people's attention or inciting them to take some action, in this case to endorse the Declaration. Only 3.4% indicated that as the main prompt for endorsing the Declaration. Lots of institutional and organizational resources go into website creation, development and maintenance, but the impact is not always so great. According to the answers received, direct contacts by the GreyNet International (almost 60%), or by a colleague (18.6%) produced the best results which should encourage us to continue maintaining personalized mailing lists and to use multiple opportunities offered by social media, such as Facebook and Twitter.

Interestingly enough, attendance the workshop where the Pisa Declaration was drafted was the reason for only 18.6% respondents to endorse it.



[6] https://goo.gl/2l8Zbx

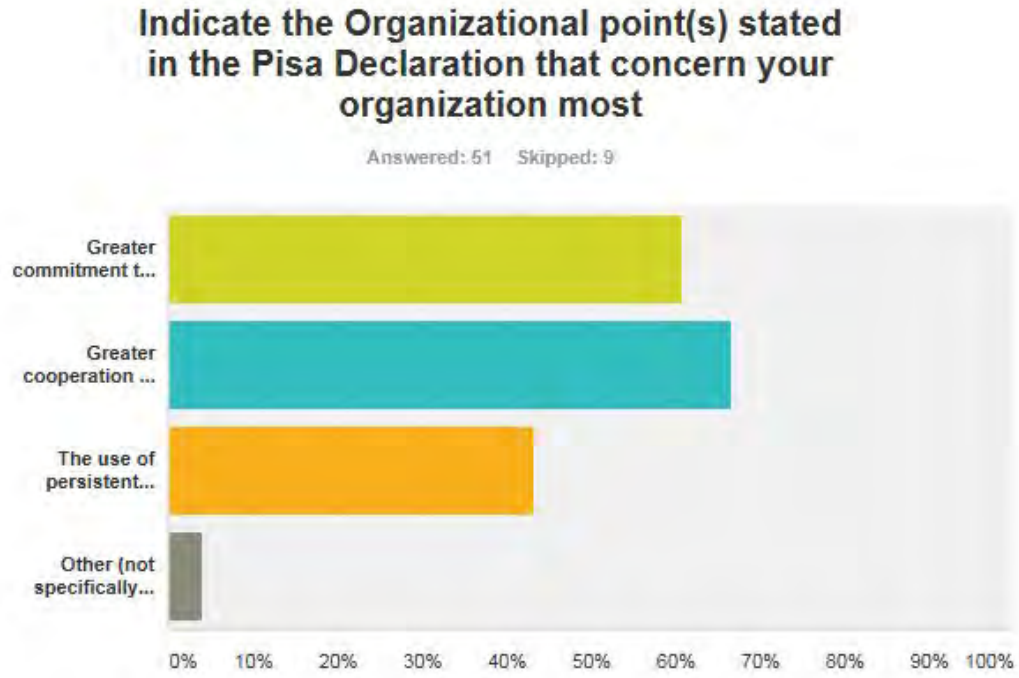
Answer Choices	Responses	
I attended the workshop where the Pisa Declaration was drafted	18.64%	11
I was made aware via the GreyGuide Portal and Repository	3.39%	2
I was alerted by GreyNet International	59.32%	35
I was alerted by acolleague	18.64%	11
Total		59

Comments (5)

Figure 2: Survey Question No. 1

Question 2: Indicate the Organizational point(s) stated in the Pisa Declaration that concern your organization most

Although the question was referring to the Pisa Declaration, it was aimed at finding out more about the main topics of interest or main concerns that organizations have regarding grey literature. Replies indicate that all three indicated areas namely open access, cooperation, and operational standards, currently represent topics of high importance and interest. In a way it is an indication of the current state of grey literature in organizations where much action is required for better processing, dissemination and use of this type of literature.

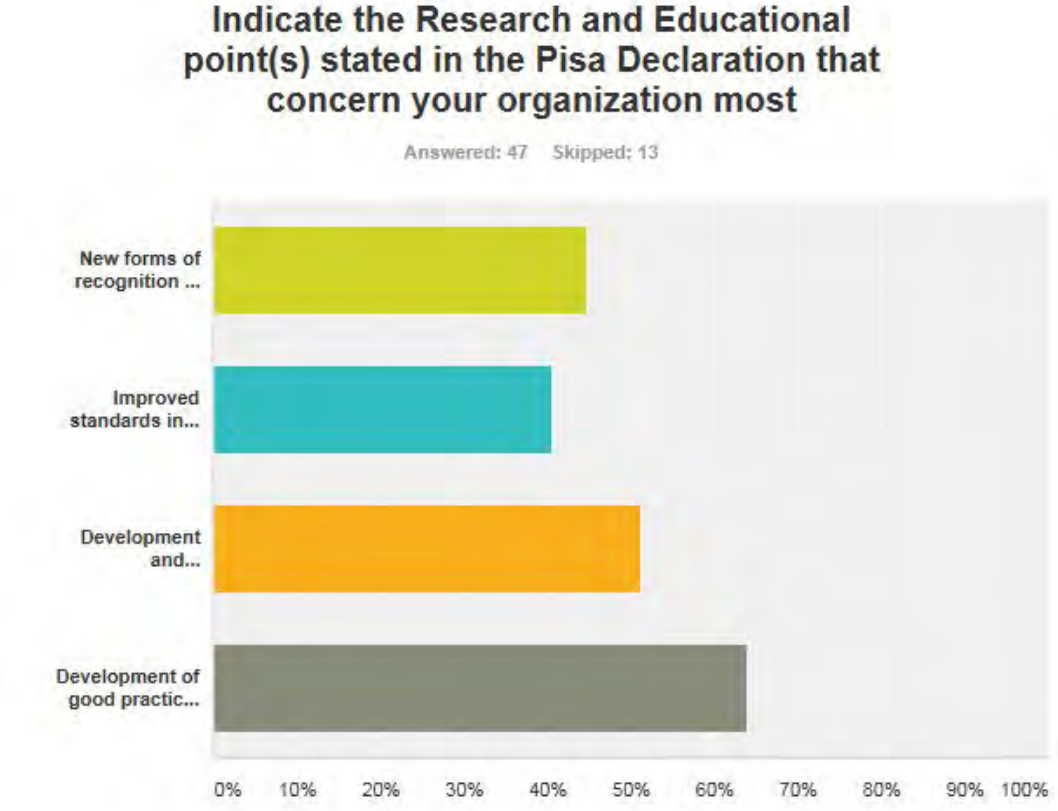


Answer Choices	Responses	
Greater commitment to open access by governments and organizations	60.78%	31
Greater cooperation and coordination among organizations engaged in the production, use, collectionand management of grey literature	66.67%	34
The use of persistent identifiers and open metadata standards for grey literature	43.14%	22
Other (not specifically mentioned in the Declaration) Responses	3.92%	2
Total Respondents: 51		

Figure 3: Survey Question No. 2

Question 3: Indicate the Research and Educational point(s) stated in the Pisa Declaration that concern your organization most

In a way question number 3 is a continuation of the previous question, since both of them try to find out more about the main topics of interest or main concerns that organizations have regarding grey literature. The difference is that this particular question is concentrated on research and educational points. Similar nature of the question and similar answers received. They also indicate that all four indicated areas – recognition, production and interoperability standards, and good practices, currently represent topics that require attention and further improvement work. Again, it is an indication of the current state of grey literature in general, with many areas and opportunities for improvements.



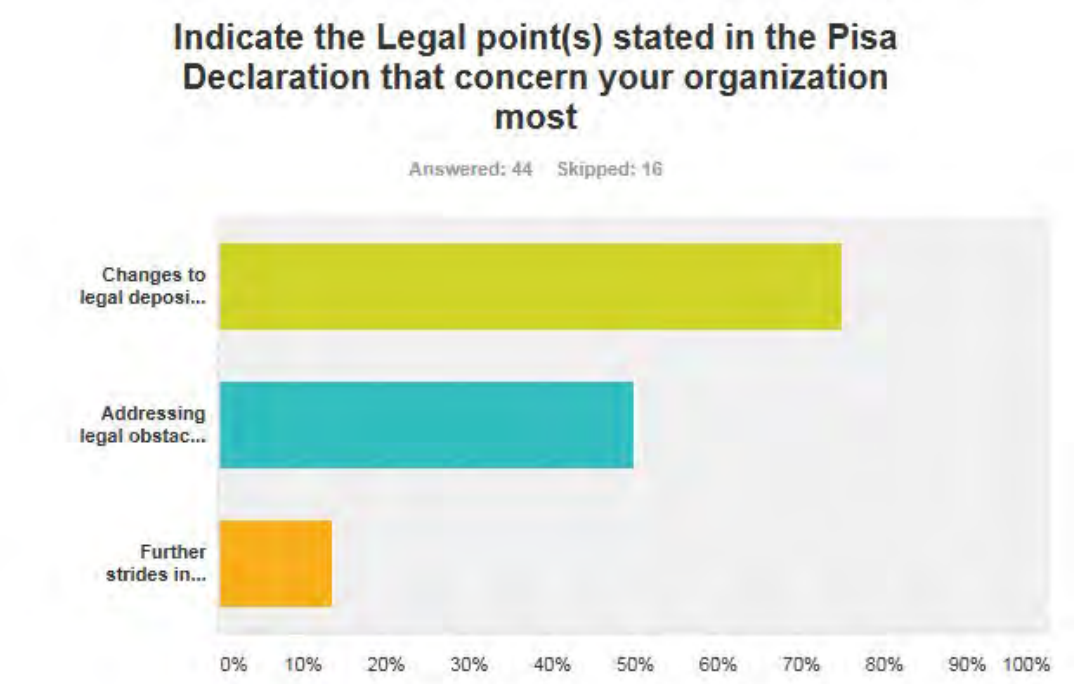
Answer Choices	Responses
New forms of recognition and reward for quality grey literature materials by governments, universities and other institutions	44.68% 21
Improved standards in the production and bibliographic control of grey literature	40.43% 19
Development and implementation of interoperable standards in the management of grey literature	51.06% 24
Development of good practice guides for the production, dissemination, and evaluation of greyliterature	63.83% 30
Total Respondents: 47	

Comments (2)

Figure 4: Survey Question No. 3

Question 4: Indicate the Legal point(s) stated in the Pisa Declaration that concern your organization most

Legal issues and protection of the intellectual property in information management and in management of grey literature is of huge concern for everyone. It is a concern to both – information providers and information users. Challenges are present on both sides, although in different forms. Information providers want to make their documents available, preferably as open source, but still protected as their unique intellectual property. Information users would like to use as much documentation, information and data as possible, but at the same time to be respectful of copyright issues. What the replies to this question on legal concerns indicate is that providers need enhanced copyright regulations that will improve the capabilities of libraries and other collecting services so that they can provide available documentation without much hindrance.



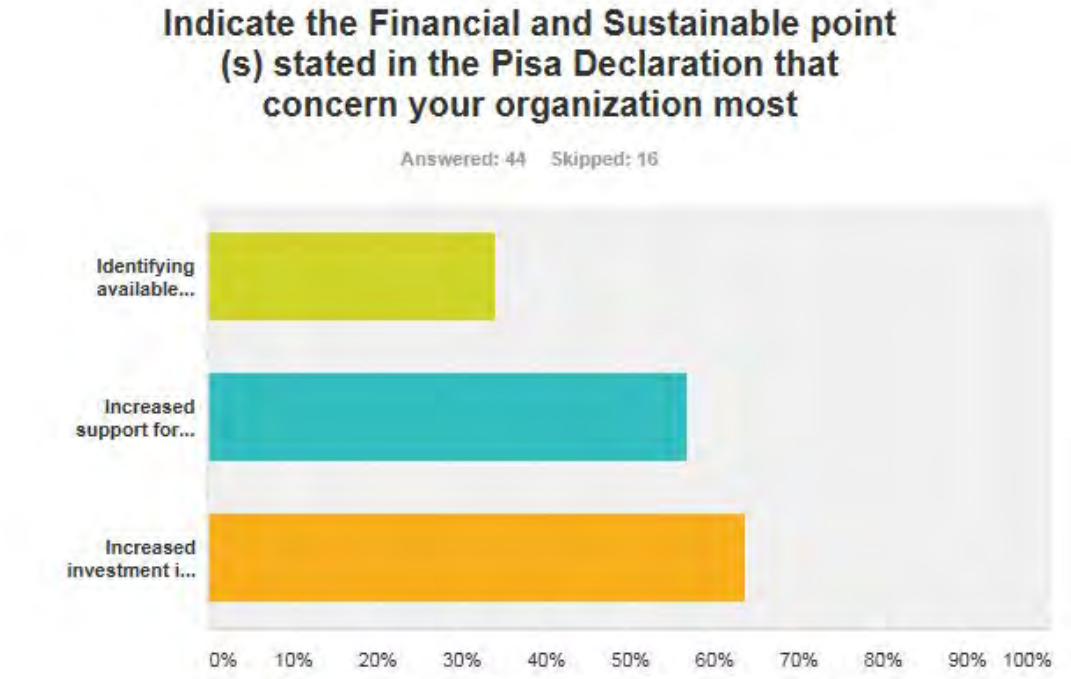
Answer Choices	Responses
Changes to legal deposit and copyright law that enhance the capacities of libraries, collecting servicesand educational institutions and programs to collect and provide access to grey literature, particularlynon-commercial public interest materials	75.00% 33
Addressing legal obstacles to the dissemination of grey literature	50.00% 22
Further strides in licensing grey content for both commercial and non-commercial purposes	13.64% 6
Total Respondents: 44	

Comments (1)

Figure 5: Survey Question No. 4

Question 5: Indicate the Financial and Sustainable point(s) stated in the Pisa Declaration that concern your organization most

Replies to this question were predictable. Grey literature, all of its sides and activities need more money, better and sustainable financing on a long run. The most urgent seems to be increased investment in infrastructure and new technologies, followed by grey literature long-term preservation. The issue of preservation is particularly vulnerable one since both, paper as well as digital collections, are disappearing quickly. At the same time users demand quick and unrestricted access to full-text documentation. This is a very huge area for further work in making grey literature more relevant and better appreciated.



Answer Choices	Responses
Identifying available funding for research involving grey literature	34.09% 15
Increased support for collection development and long term preservation of grey literature	56.82% 25
Increased investment in infrastructure and new technologies for accessing and using print and digitalgrey literature	63.64% 28
Total Respondents: 44	

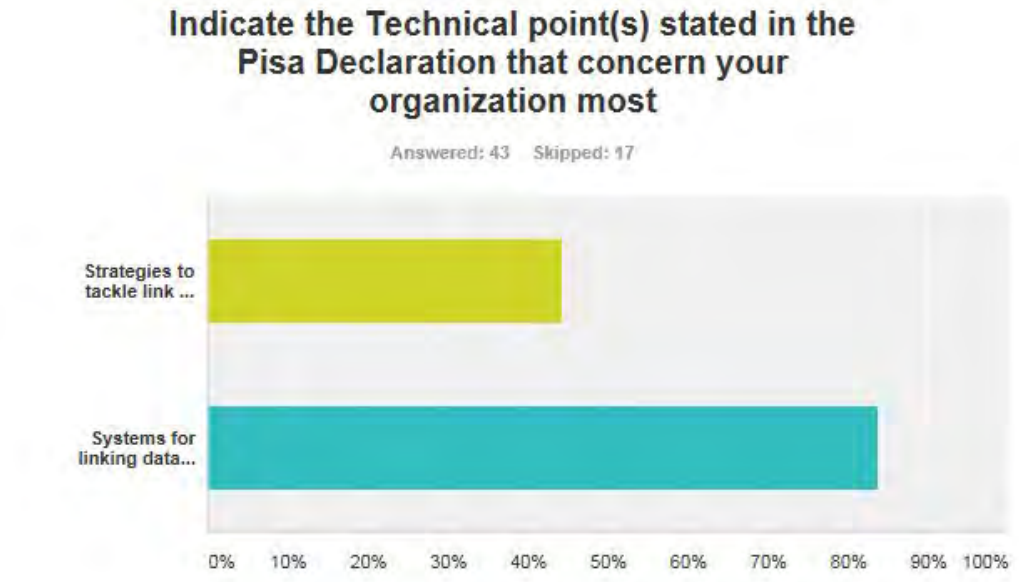
Comments (2)

Figure 6: Survey Question No. 5

Question 6: Indicate the Technical point(s) stated in the Pisa Declaration that concern your organization most

Link rot refers to problems of hyperlinks on individual websites pointing to web pages, servers or other online resources that have become permanently unavailable. This is a problem for Internet in general, not only for grey literature. The link sustainability challenge was indicated as an issue by 44% of all respondents, although we can rightfully conclude that it affects every information provider currently running or using available websites. Question of finding, repairing and preventing broken links would require further study. Use of best practices for preventing link rot, including use of DOI numbers and PURLs requires greater attention among grey literature providers.

A completely new area of linking data and other non-technical content to their grey literature publications together with interoperability standards for sharing grey literature was almost on everyone's mind. 84% of participants indicated that as the greatest technical concern for their organization.



Answer Choices	Responses
Strategies to tackle link rot and enhance the stability and accessibility of online content	44.19% 19
Systems for linking data and other non-textual content to their grey literature publications togetherwith interoperability standards for sharing grey literature	83.72% 36
Total Respondents: 43	

Comments (1)

Figure 7: Survey Question No. 6

Question 7: Is there a language(s) not listed above in which the Pisa Declaration should be translated and published? If so, please indicate here.

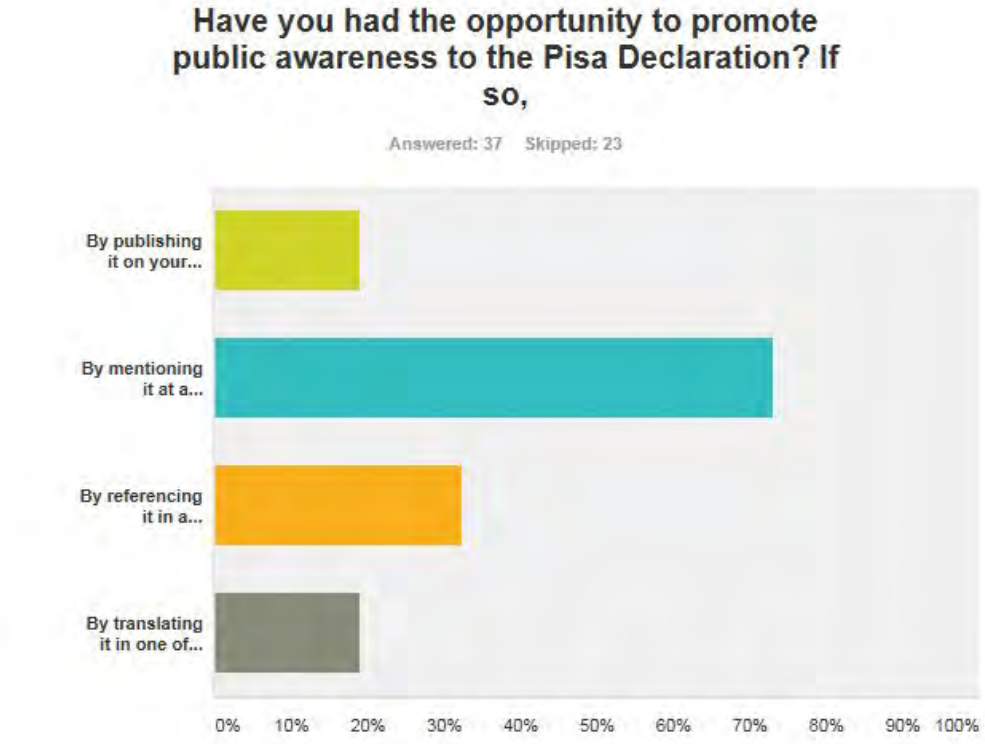
Pisa Declaration was drafted in English on 16 May, 2014. Due to wide interest and hard work of some of the members of the Grey Literature community, the Declaration was translated into 21 languages. They include: Armenian, Bulgarian, Croatian, Czech, Dutch, French, German, Greek, Hindi, Hungarian, Italian, Japanese, Korean, Macedonian, Russian, Serbian Cyrillic, Serbian Latin, Slovak, Spanish, Tagalog, and Turkish. All of the translated versions are available online[7].

The intention of the question number 7 was to find out from the respondents if there was a need for translating it into some other language. The following languages were suggested for translation and inclusion: Arabic, Chinese, Portuguese, and Korean.

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

Question 8: Have you had the opportunity to promote public awareness to the Pisa Declaration?

Starting assumptions before drafting this survey question was an impression that grey literature in general, as well as the Pisa Declaration, were not being promoted sufficiently. Replies indicate that the assumption was correct. Most of the promotion was done through ad hoc means, such as meetings or conferences. Publishing on the organizational website, as a way for providing a more sustainable presence, was exploited by few participants only. Suggestions received through comments indicated a greater need for promoting the Pisa Declaration through training and social media.



Answer Choices	Responses	
By publishing it on your organization's website	18.92%	7
By mentioning it at a conference or meeting	72.97%	27
By referencing it in a publication	32.43%	12
By translating it in one of the above mentioned languages	18.92%	7
Total Respondents: 37		

Comments (3)

Figure 8: Survey Question No. 8

Question 9: Please take a moment to record any comments, recommend-ations, amendments, or additions you consider worthwhile for the Pisa Declaration to further benefit policy development for grey literature resources.

Question number 9 asked for comments, recommendations, amendments, or additions considered to be of benefits for further policy development of grey literature. There were 12 comments received. Most of them mentioned training, development of practical manuals and appropriate standards, as well as further studies. Improving awareness among users in developing countries of grey literature benefits was also mentioned.

Question 10: Your name and email address.

It was curious to notice that out of 60 survey participants 46 identified themselves by leaving their email address. This indicates some kind of devotion to the topic being surveyed and a desire to keep in contact with colleagues and learn more about future progress and developments impacting the world of grey literature.

Concluding Remarks

The outcome of the survey leads us to conclude that direct contact via GreyNet and professionals in the grey literature community account for the majority of endorsements to the Pisa Declaration. Its placement on the GreyGuide portal was significant for its formal publication and to facilitate endorsement, however this in itself accounted for the least number of signatories. Across the board, all of the points in the Pisa Declaration are still of concern to the grey literature community.

Legal issues remain the concern not only for content and service providers but also for users of grey literature. Increased investment in new technologies enabling access to the full-text as well as related research and metadata are of equal concern. However, the need to promote public awareness to grey literature is underestimated and would contribute significantly to policy development in this field of library and information science.

Appendix 1

Pisa Declaration on Policy Development for Grey Literature Resources

May 16, 2014, Pisa

Introduction

A wealth of knowledge and information is produced by organizations, governments and industry, covering a wide range of subject areas and professional fields, not controlled by commercial publishing. These publications, data and other materials known as grey literature, are an essential resource in scholarly communication, research, and policy making for business, industry, professional practice, and civil society.

Grey literature is recognized as a key source of evidence, argument, innovation, and understanding in many disciplines including science, engineering, health, social sciences, education, the arts and humanities.
Grey literature document types in print or electronic formats include among others: research and technical reports, briefings and reviews, evaluations, working papers, conference papers, theses, and multimedia content, representing an important and valuable part of research and information.

In order to realize the benefits of research and information for scholarship, government, civil society, education and the economy, We, the signatories to this declaration, call for increased recognition of grey literature's role and value by governments, academics and all stakeholders, particularly its importance for open access to research, open science, innovation, evidence-based policy, and knowledge transfer.

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

Organizational

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

Research/Educational

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

Legal

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

Financial/Sustainable

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

Technical

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

Appendix 2

Pisa Declaration: An Assessment Study

Pisa Declaration on Policy Development for Grey Literature Resources

In the spring of 2014, a workshop took place at the Italian National Council of Research in Pisa. The topic of this event dealt with policy development for grey literature resources. Some seventy participants from nine countries took an active part in the workshop – the outcome of which produced what is today known as the Pisa Declaration.

1. How did you first come to endorse the Pisa Declaration?

- ☐ I attended the workshop where the Pisa Declaration was drafted
- ☐ I was made aware via the GreyGuide Portal and Repository
- ☐ I was alerted by GreyNet International
- ☐ I was alerted by a colleague

Other (please specify)

2. Indicate the Organizational point(s) stated in the Pisa Declaration that concern your organization most

- ☐ Greater commitment to open access by governments and organizations
- ☐ Greater cooperation and coordination among organizations engaged in the production, use, collection and management of grey literature
- ☐ The use of persistent identifiers and open metadata standards for grey literature
- ☐ Other (not specifically mentioned in the Declaration)

3. Indicate the Research and Educational point(s) stated in the Pisa Declaration that concern your organization most

- ☐ New forms of recognition and reward for quality grey literature materials by governments, universities and other institutions
- ☐ Improved standards in the production and bibliographic control of grey literature
- ☐ Development and implementation of interoperable standards in the management of grey literature
- ☐ Development of good practice guides for the production, dissemination, and evaluation of grey literature

Other (not specifically mentioned in the Declaration)

4. Indicate the Legal point(s) stated in the Pisa Declaration that concern your organization most

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

Other (not specifically mentioned in the Declaration)

5. Indicate the Financial and Sustainable point(s) stated in the Pisa Declaration that concern your organization most

- ☐ Identifying available funding for research involving grey literature
- ☐ Increased support for collection development and long term preservation of grey literature
- ☐ Increased investment in infrastructure and new technologies for accessing and using print and digital grey literature

Other (not specifically mentioned in the Declaration)

6. Indicate the Technical point(s) stated in the Pisa Declaration that concern your organization most

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

Other (not specifically mentioned in the Declaration)

Two years on, the Pisa Declaration has been translated and published in nearly twenty languages: Armenian, Bulgarian, Croatian, Czech, Dutch, French, German, Greek, Hindi, Hungarian, Japanese, Korean, Macedonian, Russian, Serbian, Spanish, Tagalog, and Turkish – all of which are online accessible via the GreyGuide Repository and Portal.

7. Is there a language(s) not listed above in which the Pisa Declaration should be translated and published? If so, please indicate here.

8. Have you had the opportunity to promote public awareness to the Pisa Declaration? If so,

- ☐ By publishing it on your organization's website
- ☐ By mentioning it at a conference or meeting
- ☐ By referencing it in a publication
- ☐ By translating it in one of the above mentioned languages

Other (please specify)

Feedback from those who endorsed the Pisa Declaration is thought to provide an up-to-date roadmap serving diverse communities involved in research, publication and the management of grey literature resources

9. Please take a moment to record any comments, recommendations, amendments, or additions you consider worthwhile for the Pisa Declaration to further benefit policy development for grey literature resources.

10. Your name and email address

A Geographical Visualization of GL Communities: A Snapshot

Gabriella Pardelli, Sara Goggi, Roberto Bartolini, Irene Russo, and Monica Monachini,
Istituto di Linguistica Computazionale "A . Zampolli", CNR Pisa, Italy

1. Introduction

> ***"Today, in the spirit of science, grey literature communities are called to demonstrate their know-how and merit to wider audiences"***
> [Farace Dominic J., 2011].

This quotation stresses the important role of the several international organizations in producing and disseminating knowledge in the field of Grey Literature (GL): the paper aims to provide a first snapshot of the geographical distribution of GL organizations and their participation to the annual International Conference on Grey Literature over the time (in the period from 2003 to 2015. See List of Conferences on Table 2).
Nowadays a visual representation of data is often associated with the traditional statistical graphs, in particular for representing complex phenomena by means of maps and diagrams, which allow a deeper and more focused analysis of the data. In our case the geographical representation of stakeholders in government, academics, business and industry aims at visualizing the GL community across the globe: it concerns 674 organizations which over the years have contributed to the development of a common vision on the most pressing issues of the field by using new paradigms such as Open Access and the social networks.

Given this scenario the GL Community is visualized by name and country of the organization and by year, as documented by the GL List of Participating Organizations published in the thirteen GL Program Books which can be found on the GreyGuide[1] site. The results are presented in the form of visual graphs, which confirm the international flavor of this field.

2. GL Community today

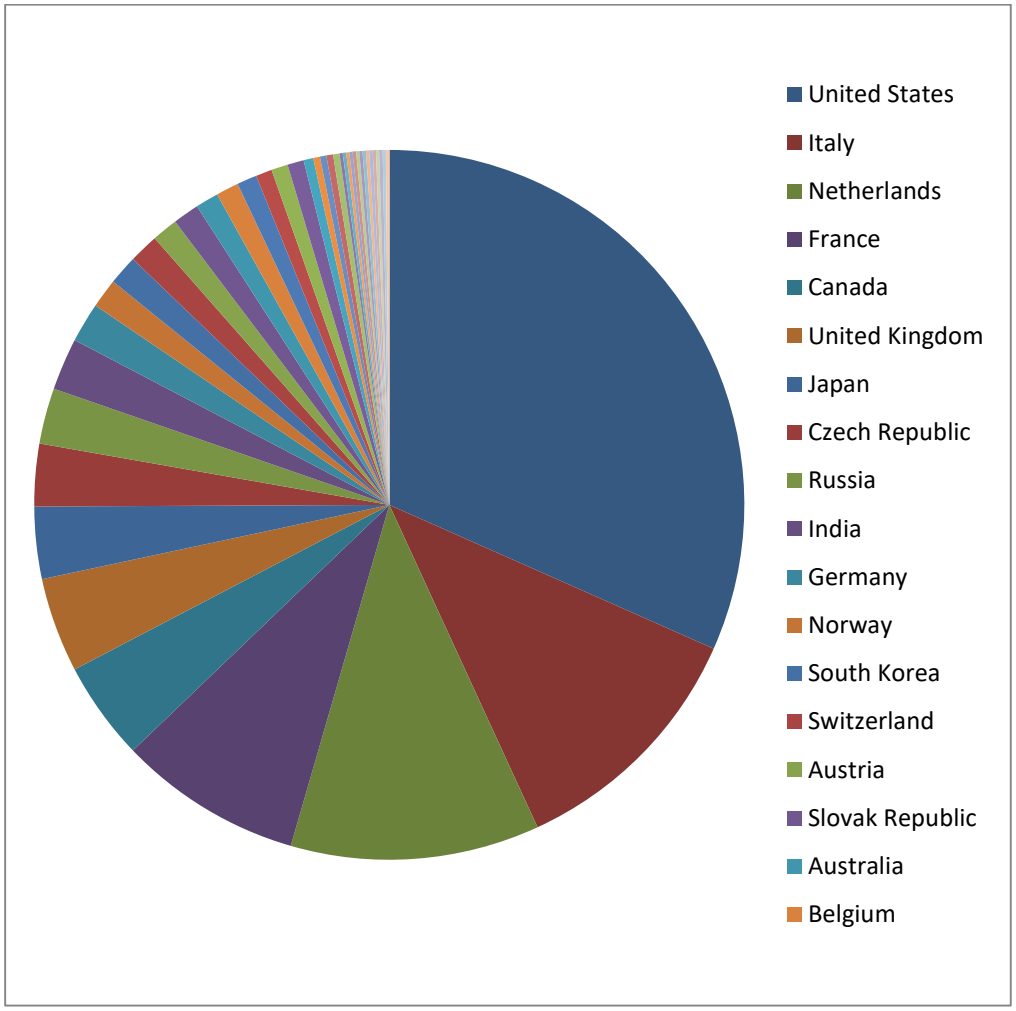
The inter-disciplinary dimension, the specialized themes and the geographical dislocation of its stakeholders are the requisites of attraction of the international Grey Literature community and these elements can represent an advantage for the whole field. Over the years universities, research centers, governmental bodies and industries presented their own research experiences, the technological solutions tested and/or adopted thus facilitating the introduction of new paradigms as well as the giving up of obsolete models.

2.1 GL Community in the world

The most remarkable figure to be reported is the substantial participation of US organizations to the GL conferences: not surprisingly the country stands at the top of the list with 216 organizations[2]. This is the chronological distribution of the American presence to the conference: 2004>32; 2005>10; 2006>25; 2007>10 2008>18; 2009>34; 2010>11; 2011>23; 2012>9; 2013>6; 2014>18; 13; 2015>13.
As shown in graphs 1 and 2, the participation of European institutions of the field is large; but over the years the community has also taken advantage of contributions from institutions coming from somehow countries such as Fiji, Finland, Gambia, Georgia, Iceland, Iran, Latvia, Luxembourg, New Zealand, Romania, Saudi Arabia, Serbia, West Indies, Zimbabwe.
Here below Table 1 shows all the countries which had at least one participating institution from GL5 to GL17.

[1] http://greyguide.isti.cnr.it/
[2] The participating organizations have been counted for each single participation (that is, they have been re-counted if attending more than one edition of the conference).

Graph 1: Visualization of GL Community in the world



Graph 2: GL Countries

COUNTRY	NUMBER OF PARTICIPATIONS
Algeria	2
Australia	7
Austria	8
Belgium	7
Brazil	2
Cameroon	1
Canada	30
Czech Republic	19
Denmark	1
Fiji	1
Finland	2
France	56
Gambia	1
Georgia	1
Germany	12
Greece	5
Iceland	1
India	16
Iran	1
Italy	77
Japan	22

COUNTRY	NUMBER OF PARTICIPATIONS
Latvia	1
Luxembourg	5
Netherlands	77
New Zealand	1
Norway	9
Poland	5
Romania	1
Russia	17
Saudi Arabia	1
Serbia	1
Slovak Republic	8
Slovenia	6
South Africa	3
South Korea	9
Spain	2
Switzerland	9
Uganda	1
United Kingdom	29
United States	212
West Indies	1
Zimbabwe	1

Table 1: Number of participations by Country.

1.	*2003 GL5 Amsterdam, "Grey Matters in the World of Networked Information"*
2.	*2004 GL6 New York, "Work on Grey in Progress"*
3.	*2005 GL7 Nancy, France "Open Access to Grey Resources"*
4.	*2006 GL8 New Orleans, "Harnessing the Power of Grey"*
5.	*2007 GL9 Antwerp, "Grey Foundations in Information Landscape*
6.	*2008 GL10 Amsterdam, "Designing the Grey Grid for Information Society"*
7.	*2009 GL11 Washington D.C., "The Grey Mosaic: Piecing It All Together"*
8.	*2010 GL12 Prague, "Transparency in Grey Literature, Grey Tech Approaches to High Tech Issues"*
9.	*2011 GL13 Washington D.C., "The Grey Circuit, From Social Networking to Wealth Creation", Library of Congress, December 5–6*
10.	*2012 GL14 Rome, Italy, "Tracking Innovation through Grey Literature", National Research Council, CNR, November 29–30*
11.	*2013 GL15 Bratislava, Slovak Republic, "The Grey Audit, A Field Assessment in Grey Literature", December 2–3*
12.	*2014 GL16 Washington D.C. "Grey Literature Lobby, Engines and Requesters for Change", December 8–9*
13.	*2015 GL17 Amsterdam, "A New Wave of Textual and Non-Textual Grey Literature", December 1–2*

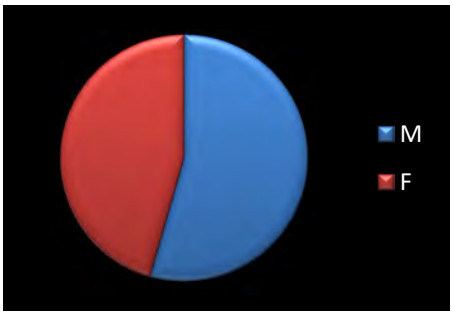
Table 2: List of GL conferences**.**

2.2 GL Community and Genre

The information about the entire set of papers presented at the GL conferences in the period 2003-2015 is available on the GreyGuide repository. In addition to the nationality of the authors, we lately decided to extrapolate the information on their gender as well: this type of analysis is usually difficult due to the various ways of writing the names (full name, initials, middle initials). It was therefore needed a cleaning process for being able to divide the authors by gender: for disambiguating the initials of the first names we used portals such as OpenGrey[3], GreyNet[4], TextRelease[5] (the section 'Who is in Grey Literature') and GreyGuide. Quite a number of authors have been identified by accessing the repository Google Scholar Citations and social networks such as LinkedIn; publishing houses online

[3] ***A*** *System for Information on Grey Literature in Europe. http://www.opengrey.eu/*
[4] *http://www.greynet.org/*
[5] *http://www.textrelease.com/*

libraries and digital archives (i.e., @rchiveSIC[6]) have been consulted as well. For example, the name 'Goggi, S.' has been easily retrieved from various repositories, disambiguated as 'Goggi, Sara' and annotated as "F" (*female)* but with other names the retrieval has been possible only thanks to the pictures available on the web (see the example of the name 'Khan, M.T.M.', identified as "M" (*male)* with the help of his picture on LinkedIn). A few unresolved cases have been annotated with a question mark. Graph 3 shows the participation by gender to GL conferences: of course the names of those who participated to more than one edition – or even presented more than one paper at the same conference - have been counted only once.
The results talk about a sort of gender balance with a slight preponderance of men: female 201, male 240.



Graph 3: GL authors' genders

3. Data Extraction and Analysis

Prior to the data analysis, a normalization process of the information provided by the authors was needed: the manual cleaning was mainly carried out on the names of those affiliations which varied both graphically and linguistically over the time. Changes concern also acronyms and abbreviations – often missing – and other types of information which help specifying the organization. Table 1 lists ten examples of variation of names.

N°	Name 1	Name 2	Variant
1	University of Pretoria – UP	University of Pretoria, UP	graphic sign (gs)
2	University of Ljubljana	University of Ljubljana, UNI-LJ	acronym (ac)
3	University of Illinois, UIUC	University of Illinois at Urbana-Champaign	information of a second level (isl)
4	Science & Technology Facilities Council, STFC	Science and Technology Facilities Council, STFC	graphic sign (gs)
5	Biblioteca Centrale "G. Marconi" CNR	Biblioteca Centrale "G. Marconi", CNR Also Biblioteca Centrale "G. Marconi"; CNR Also Consiglio Nazionale delle Ricerche, Biblioteca Centrale	graphic sign (gs)
6	Centre of Information Technologies and Systems of Executive State Authorities	Centre of Information Technologies and Systems, CITIS	information of a second level (isl) + acronym (ac)
7	Data Archiving and Networked Services, DANS	Data Archiving and Networked Services, DANS-KNAW	acronym (ac)
8	Koninklijke Nederlandse Akademie van Wetenschappen – KNAW	Royal Netherlands Academy of Arts and Sciences, KNAW	Langue (la)
9	Institute of Computational Linguistics, ILC-CNR	Istituto di Linguistica Computazionale, ILC	Langue (la)
10	University of Bergen	University of Bergen, Research Documentation Unit – UIB	information of a second level (isl)

Table 3 – Examples of names' variations

[6] *Archive Ouverte en Sciences de l'Information et de la Communication. https://archivesic.ccsd.cnrs.fr/*

Normalization has been therefore essential for the correct identification of the affiliations and the resulting calculation of their presence over the years: the same affiliations are present in different years with different authors and authors of the same institution can even present various activities at the same edition of the conference.
The assessment of the number of affiliations which participated to the GL series over this time span required their ordering in alphabetic tables and then checks on several web sites for disambiguating the graphic and linguistic variants.
Normalization was also performed on languages, for example: Nederlands Instituut voor Wetenschappelijke Informatiediensten – NIWI > Netherlands Institute for Scientific Information Services, NIWI; Consiglio Nazionale delle Ricerche – CNR > National Research Council, CNR; Istituto di Ricerche Popolazione e le Politiche Sociali, IRPPS > Institute of Research on Population and Social Policies, IRPPS.

4 Conclusion and Future Work

With this work we carried out a first introductory mapping of the researchers involved in the international community of Grey Literature, pertaining only their geographical origin. As said, this is still preliminary as other aspects would deserve attention too and thanks to the GreyGuide repository – a sort of international 'observatory' for GL – further investigations could be performed with different methodologies in order to achieve new goals.
For example, we plan to more deeply investigate the structure of the research community through the graphs of both collaboration and citation amongst authors, as a sort of social network. This process will help identifying groups of researchers who publish together or usually cite each other.
It is certainly true that the idea of visualizing the tendency of national participation to international conferences – and in our case to the GL series – over the years could be applied to similar research in grey literature thus stimulating further visual surveys from scholars in the field.

Essential References

DEL GRATTA R., FRONTINI F., MONACHINI M., PARDELLI G., RUSSO I., BARTOLINI R., KHAN F., SORIA C., CALZOLARI N. (2016). LREC as a Graph: People and Resources in a Network. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asunción Moreno, Jan Odijk, Stelios Piperidis (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation* (*LREC 2016*). European Language Resources Association (ELRA), Paris. Pages 23-28. <http://www.lrec-conf.org/proceedings/lrec2016/index.html>

DEL GRATTA R., FRONTINI F., MONACHINI M., PARDELLI G., RUSSO I., BARTOLINI R., GOGGI S., KHAN F., QUOCHI V., SORIA C., CALZOLARI N., (2015). Visualising Italian Language Resources: a Snapshot. In Cristina Bosco, Sara Tonelli, Fabio Massimo Zanzotto (eds.). *Proceedings of the Second Italian Conference on Computational Linguistics* (*CLiC-it 2015*). Accademia University Press, Trento. Pages 100-103.

FARACE DOMINIC J. (2011). Foreword. The Grey Circuit. From Social Networking to Wealth Creation. In *GL13 Conference Program and Abstracts*: Thirteenth International Conference on Grey Literature . GreyNet, Grey Literature Network Service. – Amsterdam : TextRelease, 2011. Page 3.

ZEQIAN S., OGAWA M., SOON TEE TEOH, KWAN-LIU MA (2006). BiblioViz: a system for visualizing bibliography information. In *Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation - Volume 60* (APVis '06), Kazuo Misue, Kozo Sugiyama, and Jiro Tanaka (Eds.), Vol. 60. Australian Computer Society, Inc., Darlinghurst, Australia, Australia. Pages 93-102.

Teaching and Learning about Grey Literature

Results from a Poster Presented at the 18th Grey Literature Conference

Lynne M. Rudasill,
Center for Global Studies,
University of Illinois at Urbana-Champaign, United States

Formal opportunities that provide teaching and learning related to the creation, dissemination and conservation of grey literature are few outside of workshops provided by the GreyNet organization. Research by Rabina revealed that library school students in the U.S. received most of their training in the use of grey literature on the job rather than as a part of a designated course. Further, she posits that a cross-curricular approach would most closely reflect the broad scope of grey literature should it be more prominently placed in the curriculum of library and information programs. (Rabina, 2011) With this in mind, the author proposed a poster session on the topic of how we can improve understanding of the creation, access, and preservation of grey literature in I-Schools and in general. The poster included four related questions:

- How do we improve understanding through formal learning?
- How do we expand the general audience and participation?
- What are our learning objectives?
- What methods should we use to disseminate information on grey literature more broadly?

We then provided "sticky notes" to participants and invited them to provide us with their comments. Participation was quite robust with approximately twenty-five individuals participating in discussions about the topic and providing notes on their ideas. The results of the conversations and notes can be broken down into the following topic areas which have been ranked by times mentioned.

Results

Most of the comments related directly to paths to promote grey literature more broadly. Most frequently mentioned were freely available webinars to provide outreach and training. In a closely related area, promotion of grey literature through video tutorials and other social media was also seen as high priority. A number of individuals were surprised that a robust listserv was not available for the discussion of grey literature and surrounding issues and suggested that one be developed.

A second area of major concern was the preservation and indexing of grey literature. Several individuals strongly supported the concept of better search capabilities for grey literature through development of better metadata for these items. These discussions also included suggestion of the development of a Wikipedia type of open access resource for the storage of otherwise non-indexed articles, reports, etc. Related to this was the promotion of case studies using grey literature and the creation of a bibliography for newcomers.

Finally, there was interest in incorporating more information on grey literature in library and I-School courses and also bringing this information into the undergraduate classroom. We should also be sure that our scholars are aware of this type of literature. Mention was also made of the importance of introducing public library users to grey lit as well as providing information on FOIA laws. A variety of other suggestions were made to improve the "branding" of grey literature, develop alliances with other professional organizations, and perhaps even develop professional degree related to grey literature.

Conclusion

The GreyNet LIS Education and Training Committee headed by Marcus Vaska will be actively working on teaching and learning for grey literature in the coming year. Please let us know if you have any other suggestions to raise the profile and use of grey literature.

Reference
Rabina, D. L. (2011). Course and Learning Objective in the Teaching of Grey Literature: The Role of Library and Information Science Education. Grey Journal (TGJ), 7(3), 160-165

Appendix – Transcript of the results from the conference poster:

There should be webinars for users on the use and value of GL

Quarterly webinars like "Help I'm an accidental government librarian!" on various topics

- Massive indexing effort
- Build researcher/student awareness
- Digitization efforts (ex. See NYPL example) also as institutions create grey lit, they should archive PDF copies

- Hands on training
- Concentrate on
  - Collection development
  - Users/customers
- Offer professional degree/status
- Accreditation? Why not
- Incorporate gl repositories in bigger search engines

Depositing GreyLit in a repository for discovery and preservation and harvesting to aggregators

What is the difference between GL & oA Users better understand the latter.

We need a listserv!!
Teach grey literature in LIS programs

"Streaming" grey literature

1. Use real life inquiries
2. Reach out to public libraries
3. How do I find grey literature in health and medicine
4. Free workshops in the PL and other LIS mtgs,

Possibly add a listserv? (how can these standards be disseminated?)
Provide a bibliography for newcomers?

Why reinforce GL – instead promote metadata & more comprehensive access.

How to evaluate the authoritativeness of information contained in GL. Does GL have an advantage (e.g. timeliness) over traditional professional publications?

What is the brand?
When talking to different communities, how do we describe what grey lit is?

Improve researcher understanding of locating grey literature.
- Training
- Sessions.

The term GL is confusing to many professionals, the terminology using data vocabulary is more attractive.
Alliance w/ other professional organizations?

Case studies

Presenting grey literature on social media platforms

YouTube

Make a video narrated with simple stick figures and examples

It is not important for end users to know what is GL. For them is important to get document, dataset, etc. withouy the knowledge that is GL. Let's do the GL more accessible = store it, preserve it and make it searchable.

The best way to expand our audience is to push librarians and faculty at universities to force students to use grey literature. It also be influenced in undergrad library classes so that students can use the searching skills their entire college experience.

- Share use cases
- Promote open use of discussion lists
- Trainings on transparency/FOIA law + the like

Provide more broad search
Capabilities i.e., websites, Google
Outside normal channels

If institutions and individuals are storing data sets why is there not an open source that users can upload information on topics. Such as a Wikipedia for non-indexed articles.

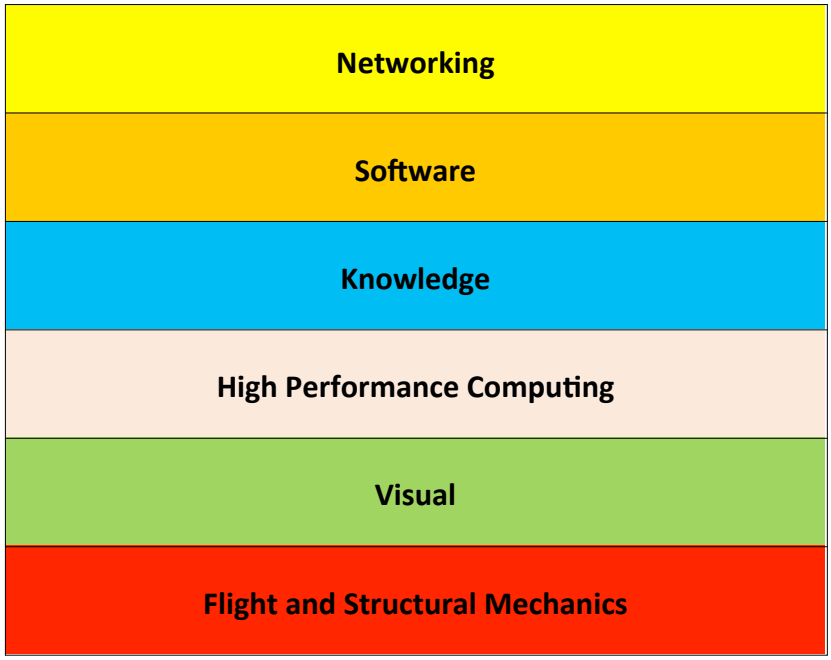
Library Schools
Need to start with the library school administrators who can expand their program curriculum

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

an Institute of the National Research Council of Italy CNR

ISTI is committed to produce scientific excellence and to play an active role in technology transfer.
The domain of competence covers Computer Science & Technologies and a wide range of applications.

The research and development activity of the Institute can be classified into 6 thematic areas



CNR-ISTI, Via G. Moruzzi 1
56124 Pisa (PI), Italy
Area della Richerca del CNR

Contact: +39 050 315 2403
segreteriascientifica@isti.cnr.it
http://www.isti.cnr.it

A terminological "journey" in the Grey Literature domain

Roberto Bartolini, Gabriella Pardelli, Sara Goggi,
CNR, Istituto di Linguistica Computazionale, "Antonio Zampolli"
Silvia Giannini and Stefania Biagioni,
CNR, Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", Italy

1. Introduction

"It is by means of terms that the expert usually transfer their knowledge and again through terms scientific communication reaches the highest effectiveness. Therefore we can assert that terminology - in the sense of a set of representative and domain-specific units - is necessary for representing and connecting specialized fields as well as any attempt to represent and/or transfer scientific knowledge requires, more or less extensively, the use of terminology." (Cabré, 2000). "When we read the articles or papers of a particular domain, we can recognize some lexical items in the texts as technical terms. In a domain where new knowledge is generated, new terms are constantly created to fulfill the needs of the domain, while others become obsolete. In addition, existing terms may undergo changes of meaning…" (Kageura K., 1998/1999).

Specialized lexicons are made up of the terms which are specific to each field of knowledge, «a subset which is distinct but not separated from the common language» (Cassese, 1992): it is usually difficult to extract the relevant domain-specific terminology, meaning to discern terms which belong to a specialized glossary from those belonging to the common dictionary.

The interest in the study of terminology and the "truth" contained in the above definitions has led us to make a "journey" in the Grey Literature (GL) domain in order to offer an overall vision on the terms used and the links between them.

Within this scenario, the work analyzes a corpus constituted of the entire amount of full research papers published in the GL conference series over a time-span of more than one decade (2003-2014) with the aim of creating a terminological map of relevant words in the various GL research topics. "… corpora used to extract terminological units can be further investigated to find semantic and conceptual information on terms or to represent conceptual relationships between terms. (Bourigault D. et al., 2001). Another interesting inquiry is the terminology used in the GL conferences for describing the types of documents which can be detected (Pejšová P. et al., 2012).

2. GL Corpus and method

The work is split up in four sections: creation of the corpus by acquiring the digital papers of GL conference proceedings (GL5 – GL16)[1]; data cleaning; data processing using the described NLP pipeline; terminological analysis and comparison. The corpus - made up of 231 research papers (for a total amount of 785.042 tokens) - was processed using a Natural Language Processing (NLP) tool for term extraction developed at the Institute of Computational Linguistics "Antonio Zampolli" of CNR[2] (Goggi et al. 2015; 2016).

This tool is what is called a "pipeline" - that is, a sequence of different tools - which extracts lexical knowledge from texts: in short, this is a rule-based system tool for knowledge extraction and document indexing that combines Natural Language Processing (NLP) technologies for term extraction.

The NLP pipeline analyzes textual data thanks to generic tools and its result is an annotated text that allows for terminological extraction of relevant concepts.

More in details, these are the steps which it follows:

- transformation of the original document, in our case in Word format, in plain utf-8 format text;
- use of some pre-existing software tools for:
  1. sentence splitting: dividing the text into sentences
  2. word tokenization: splitting sentences into words

[1] Kindly authorised from Greynet International, http://www.greynet.org/.
[2] CNR stands for National Research Council, Italy, https://www.cnr.it/.

3. lemmatization and morphological analysis (part of speech tagging)
4. basic syntactic analysis (chunking: dividing the sentence into non recursive constituents)
5. parsing with the **Ideal** dependency parser, a rule-based system whose specific rules were developed for both Italian and English. This tool was developed specifically for the MAPS project, being the most important part of the NLP pipeline.

The output of the chunking phase produces an intermediate annotated document preliminary to terminology extraction performed by the **Ideal** parser, which relies on rules, written in an ad-hoc language, designed to extract all simple and complex noun phrases in the text[3].
Terminological extraction is in turn necessary in order to be able to correctly index the document in the document base to be later semantically searched. The **Ideal** extraction tool takes the chunked text as input, containing all the required morpho-syntactic information.
The output of this terminology extraction pipeline is a set of terms in a standardised *Jason* format.
Within our corpus made of GL articles, this NLP tool – already used as semantic engine within the MAPS project (GL16 and GL17 papers) - extracts a list of single (monograms) and multi-word terms (bigrams and trigrams) ordered by frequency with respect to the context.

3. Terminological analysis

The terminological analysis started with the identification of the monograms of high, medium and low frequency within the glossaries provided by the extraction. This first step gave us an overview of single-terms used in the papers. The study of the terms grouped according to their frequency allowed us to: a) select some of the terms most frequently used; b) examine their co-occurrences; c) determine the variations between them. We continued the terminological analysis with the observation of fragments of taxonomic chains in order to shed light on the usage of specific terms within the topics of the various GL conferences. Through these steps it was possible to monitor the terminological flow and to indicate the resulting lexical trend within the GL domain.

3.1 High, medium and low frequency

For frequency segment of vocabularies we mean the organization of words by decreasing frequencies, starting from the word with $freq_{max}$ and coming to those with $freq_{min}$, usually with only one occurrence (hapax). The occurrences can then be divided into three groups (high, medium and low frequency): in the high segment each word has a different number of occurrences, the limit between the high and medium frequencies being placed immediately above the first parity, that is, the first pair of words that occur the same number of times. To determine the $freq_{min}$ segment and separate it from the mid-range, it is necessary to start from the bottom, i.e. from the hapax, and consider the first gap in the consecutive number of increasing occurrences. After having organized the terms it results that the highest percentage of terms is to be found in the lowest frequency segment: this applies to all GLs'. The GL16 and GL6 glossaries stand out for the substantial amount of terms in the highest segment while the medium segment can be allocated to GL5 followed by GL14.

In Table 1 and Table 2 (Appendix 1) we grouped, respectively, the terms of the highest and medium segment of each GL corpus. The following categories have been excluded from the visualization: adjectives with a semantically low relevance with respect to the context (such as "new", "coastal", "public", etc.); acronyms and generic nomenclatures of bodies, proper names of individuals and institutions.
It was not possible to representing in a table the data with a low frequency given their huge extension; however a section of the lexicon of this segment has been analyzed because it was considered as much relevant.

[3] The extractor works on the chunked text searching for patterns such as nominal phrases (monograms) and nominal phrases followed by one or more adjectival or prepositional phrases (bigrams and trigrams).

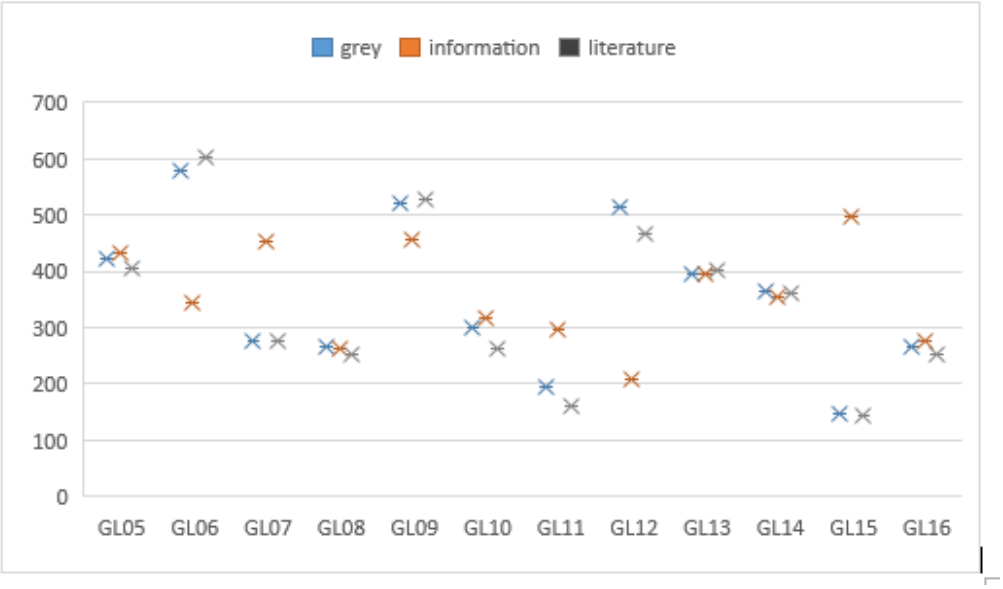
In Appendix 1, we can read the terms occurring most frequently in the high segment: the only two monograms which consistently remain in this segment and in all GL glossaries are “Literature” and “Research”.

We retrieved words such as “information” or “document” that have a very wide semantic content as well as words closely connected with the specific domain of Grey Literature such as “literature” and “grey”. There are also some terms linked to specific documentary categories such as “report”, “journal” and “thesis”.

Although “Information” is the monogram with the highest frequency in the entire corpus (4302 occurrences) it occupies the second place in the table representing the high segment terms : the first position belongs to “literature”, one of the more content-related terms, while “grey” shows 3851 occurrences in the high segment out of a total of 4298 in the whole corpus.
“Grey” appears as monogram also in the following forms: “e-grey”, “metagrey”, “non-grey”, “opengrey”. The acronym “GL” occurs 1025 times in the entire corpus; the word “information” appears as monograms also in the following form: “Bioinformation”, “Cs-Information”, “Cultural/Information”, “Data/Information”, “Librarians/Information”, “Library-Information”, “Meta-Information”, “Misinformation”, “Novel-Information”; and “literature” appears as monogram also in the following forms: “grey-literature-typology” and “sub-literature”.

3.2 Mapping

We started the terminological mapping from observing the term that occurs most frequently in the entire corpus: “information” and the two terms more closely related to the context, "grey" and "literature".
Graph 1 shows that the terms “grey” and “literature” have the highest frequency in GL6 (2004) and the lowest in GL15 (2013) while the term “information” has the peak in GL15 (2013) and the bottom in GL12 (2010).



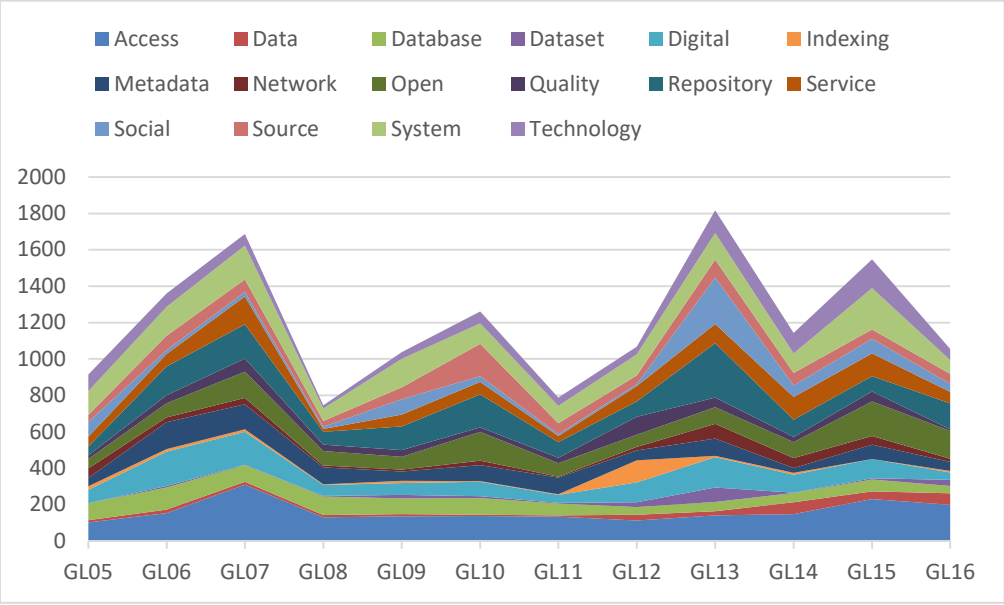
Graph 1 – “Grey”, “Information”, “Literature” – Trend over the years

As expected, the bigram “grey literature” is the most used with 2816 frequencies in the entire corpus while the bigrams “grey material” (66 occurrences) and “grey document” (98 occurrences) are not present in all GL proceedings and their frequencies are much lower. The bigram “grey documentation” only appears in GL5, GL9 and GL16. Among the other bigrams we find: “grey medical”, “grey document”, “digital grey”, “grey publisher”, “grey content”, “grey object”, “grey resource”, “grey collection”. Amongst the trigrams we have: “grey literature collection”, “grey medical literature”, “grey literature community”, “grey literature document”, “grey literature repository”, “grey literature resource”, “grey literature material”, “grey literature typology”, “grey literature report”, “digital grey literature”, “grey literature field”, “grey literature problem”.

In addition to the pair "grey literature" the term "literature" appears in the following bigrams: "medical literature", "conventional literature", "type literature", "literature repository", "literature collection", "repository literature", "use literature", "access literature", "journal literature", "literature collection", "conference literature", "trade literature", "definition literature", "literature document", "topic literature", "literature repository", "literature review". The trigram "non conventional literature" is only used in GL7 and GL14 terminology. Excluding the already mentioned trigrams in which "literature" appears associated with grey, there are: "bibliographic control literature", "new generation literature", "scholar information literature", "digital curation literature", "digital repository literature", "strategic collection literature", "literature network service", "web-based dissemination literature", "digital library literature".
The most common bigrams with the term "information" are in GL15: "Information object" is the top term (39 occurrences) while the bottom one is "Information retrieval" (17 occurrences) in GL14. Amongst the others we find: "information system", "source information", "scientific information", "information interaction", "information system", "internet information", "access information".

Looking at trigrams, "Open Source Information" is the top term with 228 occurrences and "Heterogeneous Information Object" the bottom one with 56 occurrences. Others are: "Research Information System", "Information Distribution System", "Public Health Information", "Source Information Product", "Carbon Dioxide Information", "Grey Literature Information", "Scientific Information System".
All the given lists of terms are ordered by descending frequencies.

Hereafter the analysis focused on some terms traceable in the three segments: given the dimension of the corpus and the long time-span taken into exam, the terms have been chosen according to their technical connotation with respect both to the context where they are placed and to a very dynamic and cross field, Information and Communication Technology (ICT): "access", "data", "database", "dataset", "digital", "indexing", "metadata", "network", "open", "quality", "repository", "service", "social", "source", "system", "technology".

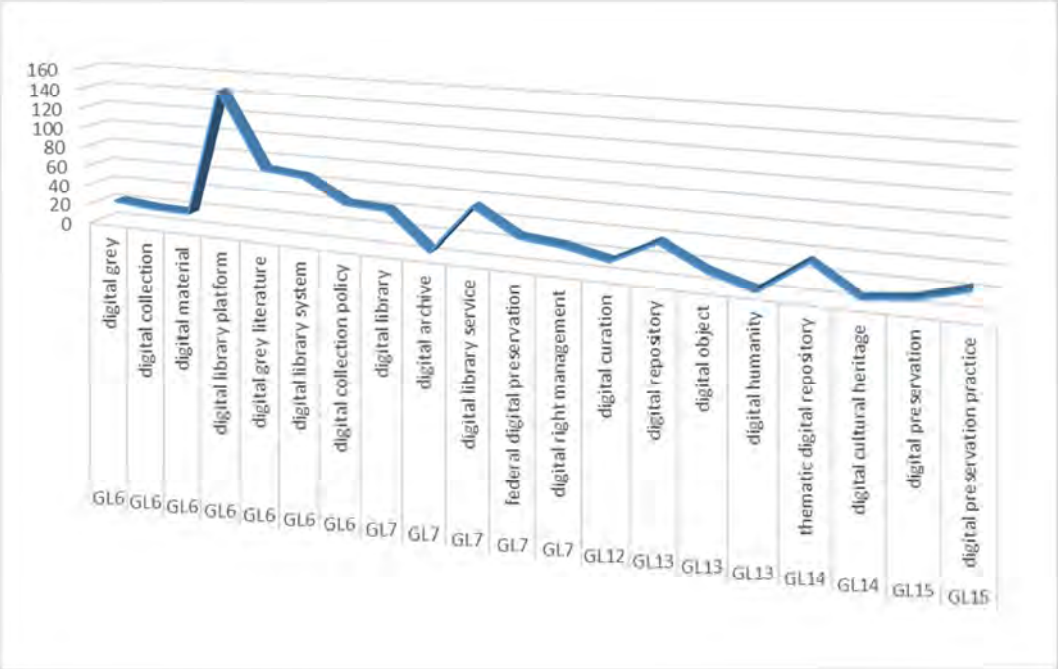


Graph 2. – Selected terms

Graph 2 shows the trend of the selected terms over the years: it is clear that - with the exemption of "indexing" and "dataset" – all of them are occurring in each GL glossary. Generally, there are monograms which seem to be constantly used and therefore their trend over the time is stable (e.g. access, database and digital) while the vast majority of terms alternate high and low frequency peaks.

The monogram “access” has the highest number of occurrences (1928) and “dataset” the lowest (196); amongst the most frequent terms, also “system”, “repository”, “open” and “digital” can be found.

Let us start our investigation from one of the most versatile adjectives of the corpus: “digital”.
Graph 3 shows the bigrams and trigrams this term form with the several nouns: “digital library” and “digital library platform” are the most recurring Multi-Word Expressions (MWE). The overview provided by the list of selected terms also points out some nouns and verbs which combined with the adjective “digital” - though with relatively low frequencies - disclose the technological nature of the GL community: infrastructure, platform, system, software, network. The MWE “digital humanity” and “cultural heritage” represent entire branches of knowledge whose activities require an expertise crossing from computer science to social and human sciences.
Among bigrams: “digital library” appears in 2005 (GL7). The community does not neglect relevant contents such as “digital preservation” which appears in 2013 (GL15) and even uses the trigram “digital preservation practice”. Among trigrams: “digital library platform” has the highest frequency in 2004 (GL6). In most recent years (from 2013 onwards) s such as “digital repository” and “thematic digital repository” replace others like “digital library” and “digital library service” thus revealing new demands for identification, accessibility, interoperability and reuse of the scientific data they host, as well as the need of ad-hoc services for those specific contents.



Graph 3. “Digital” – bigrams & trigrams

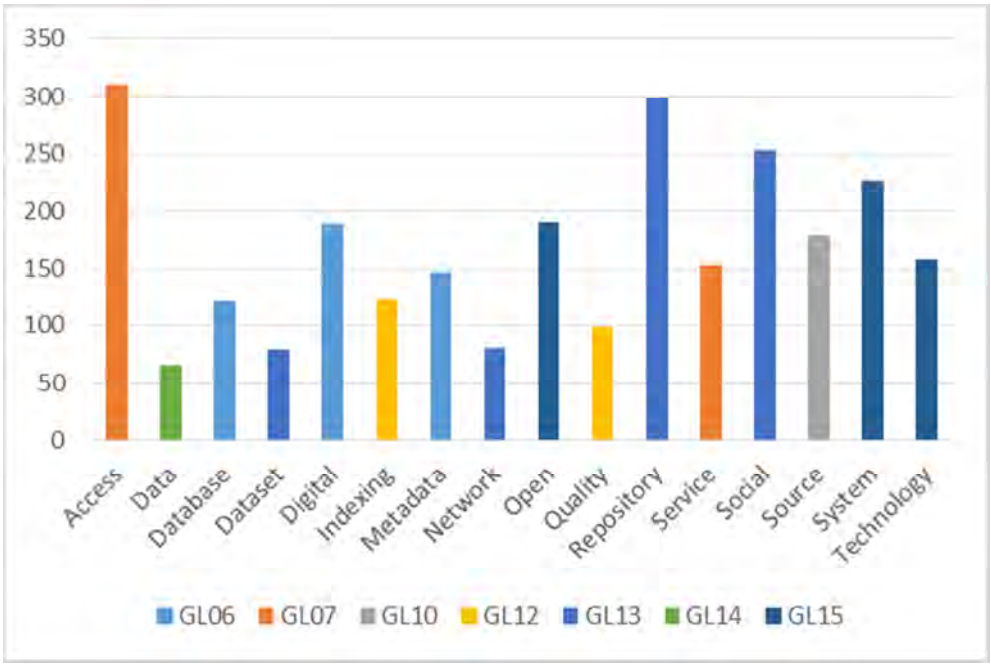
The term “data” shows the highest frequency as a bigram, “big data”: it introduces the set of problems about gathering, managing, representing and accessing huge volumes of data which are dynamically generated from various sources. The bigram first appears in GL14 (2012) and then again the next year thus witnessing the community’s immediate appeal for the subject; as a trigram is mostly in combination with terms like “discovery”, “service” and “product”.
The term “database” cannot be neglected too: it is used and reused in different contexts as a synonym of an archive of structured and connected data and occurs in the entire time-span associated with various semantic values: "citation databases", "technological databases", "grey literature database". “Database” and “metadata” register the highest number of occurrences in the papers of the GL6 (2004) conference, exactly like “digital”.
The exam of the term “metadata” points out its presence in all GLs in the mid-range segment and already in GL6 (2004) and GL7 (2005) in the high frequency one: these are years when there is a considerable discussion over themes as standards, document

management and fast information retrieval systems and the term is often found in association with nouns and adjectives which highlight the importance of properly describing and organizing documentation in the field of digital libraries: “metadata schema”, “navigational metadata”, “descriptive metadata”, “metadata format”, “metadata harvesting” and again “Dublin Core metadata” (the highest percentage), “right management metadata”, “standardized metadata schema” and “metadata quality control”. In GL7 and GL10 “metadata” is often combined with standards and schemes such as Dublin Core and Cerif.
The term “dataset”, in the two variations “dataset” and “data set”, appears in 2005 and remains constantly present in the following editions forming the most frequent bigrams “scientific dataset” and “dataset archive” while is more occasionally associated with “accessibility”, “collection” and “management”.
Already in GL6 (2004) the GL community faces the need to examine the quality of information available on the web: the term “quality” is repeatedly associated with “assessment” or “control”, in particular in the forms “metadata quality control”, “quality assessment metadata”, “quality information”, “quality performance”, “high-quality information” and “metadata quality certification”.
Another interesting term is the adjective “social”. Although we found the topic “Social Networking” only in GL13 (2011), this bigram is in use since GL7 (2005) and the monogram “social” is steadily used in the GL lexicon since GL5 (2003). The adjective “social” is combined with a large number of nouns to form bigrams, trigrams and strings of words with a strong semantic impact. In GL8 the multi-word expression "social network" appears, as a “neologism”, in the GL lexicon. Other linguistic forms emerging from the terminology are linked to the same concept: "virtual social networking”, "social networking tools", "social networking sites", "new social networking technologies". The MWE “social media” was “born” instead in the GL9 conference (2011).

The bigram “open access” which represents one of the most studied research fields in recent years, is a constant feature in the grey literature lexicon. It is in fact used since the far GL5 (2003) in the two graphic variations “open access” and “open-access” that coexist in some GLs’. From the separate analysis of the bigrams formed by “open” and “access” it can be noted that the most frequent is anyway the one which combines them; the monogram “access” then constitutes other bigrams (amongst the others “access information”, “access literature” and “access model”) and trigrams, once again with “open”: “open access model”, “open access repository” and “open access movement”. In order to avoid “open” from the lexical forms taken into exam, the lowest frequencies should be analyzed for finding forms like “sustained access information”, “access datum repository”, “public access resource”. But “open” often creates MWE also with other terms: “open archive”, “open source”, “open repository”, “open model”.
In our context the term “technology” is related to telematics and computer science applications to the documentary field: the single term is paired with “information technology” while the trigram is "technology information system". Information management is represented by nouns such as “system” and “source”: both words are also retrieved in the lexical forms "information system ", "information system database", "electronic sources", "open source repository". A special case is the word “service” which is very frequently used for defining activities for the users of the Internet: “information services", "integrative web services".

3.3 GL Conference topics

The flow of themes discussed in these years at GL conferences is represented by the topics appearing in the twelve Call for Papers (Appendix 2).
Therefore the previous selected terms have been analyzed in relation to the topics of all GL conferences by retrieving the frequency peaks of the chosen terms and then verifying when they occurred.



Graph. 4 – Terms and Topics

From Graph 4 it is clear that the peaks of frequency are limited to certain years: 2004, 2005, 2008, 2009, 2010, 2011, 2012 while the other editions are lacking. The highest frequencies occur in GL7 with the term "access" and in GL13 with "repository" and "social. The word "repository" is never found amongst the topics in its singular form but rather diffusely as "repositories" since GL6 (2004) and then again in 2005, 2006, 2008 and 2009 combined with "collection", "metadata" and "grey literature" for creating "Institutional Repository "and "Grey Literature repository". Again "repository", together with "dataset", "network" and the already mentioned "social", counts the highest number of occurrences in the GL13 papers, where some of the topics were "Social Networking", "Special Collections", "Open Access and Wealth Creation", "Data Frontiers".

The maximum number of occurrences of the terms "digital", "database" and "metadata" dates back to the GL6 (2004) conference which introduced the following topics: "Institutional Repositories", "Use Analysis", "IT & Research Initiative", "Knowledge Management and Dissemination", "Collection Development and Resource Discovery".

In the same year the adjective "digital" registers the highest frequencies with the two forms "digital library" and "digital library platform". It is curious to note that the bigram "digital library" never appears amongst the GLs' topics notwithstanding it is widespread within the articles and, even more curios, the monogram "digital" is never used either. The same for "database" while "metadata" appears only once, in the GL8 Call for Papers.

In GL14 (2012) "data" and "indexing" register their peaks: in this year the chosen conference topics were "Tracing the Research Life Cycle", "Tracking Methods for Grey Literature", "Adapting New Technologies", "Repurposing Grey Literature".

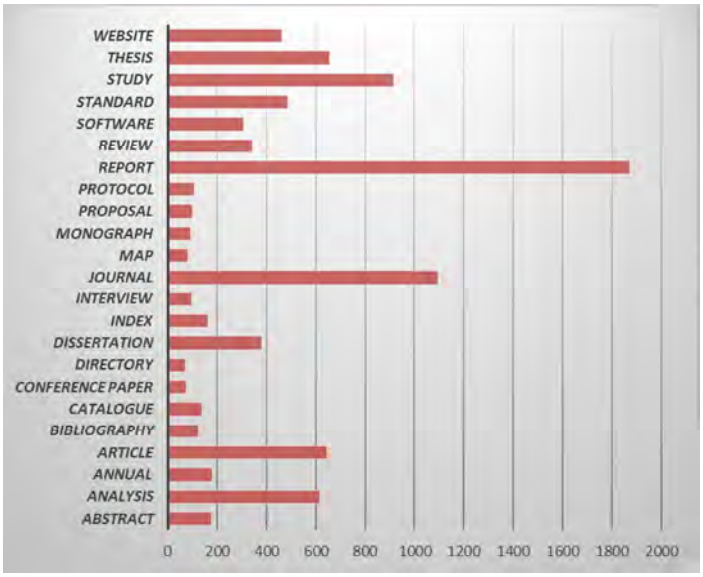
Finally three topics are dedicated to "open access" in GL conferences: "Open Access to Grey Resources", "Open Access and Wealth Creation" and "Open Access to Research Data" (GL16 - 2014).

3.4 Types of documents

This last chapter is dedicated to the terminology used for describing the types of documents occurring in the corpus.

The analysis of terminology adopted for describing the types of documents started from the entries of the *Vocabulary of the types of Grey Literature (*2011) which has been considered as the reference model. It is though important to take into account the possibility that the terms extracted from the corpus do not necessarily describe the type of GL documents because it was not possible to verify automatically the actual correspondence between the term and its context. An outstanding example is "journal" which can easily refer to the title of a publication.

From this perspective, the presence of the *Vocabulary* terminology within our corpus has been verified: the table in Appendix 3 lists the terms appearing in the various GLs and their quantitative consistency. This table is ordered by frequency and the results – in terms of the most occurring terms - are therefore very clear.
In the attempt of making a partition of this list – however arbitrary - we can circumscribe a first area where the frequencies decrease from 1871 to 307 and the terms retrieved are: "report", "journal, "study", "thesis", "article", "analysis", "standard, "website, "dissertation, "review", "software". In the intermediate area where frequencies decrease from 196 to 30 (with a remarkable gap between the last occurrence of the first zone and the first occurrence of the last zone) the terms found are the following: "dataset", "annual", "abstract", "questionnaire", "index", "patent", "catalogue", "bibliography", "annual report", "protocol", "proposal", "interview", "monograph", "map", "conference paper", "preprint", "directory", "newsletter", "manual", "bulletin", "curriculum", "poster".
The terms used with the lowest frequency (from 24 to 1) for describing the types of documents are: "brochure", "proceedings", "government document", "glossary", "memorandum", "handbook", "timeline", "announcement", "conference program", "essay", "press release", "chronicle", "leaflet", "course material", "informative material", "normative document", "anthology", "research plan", "syllabus", "tertiary source", "corporate literature", "habilitation thesis", "image material", "legal document", "guidebook", "technical documentation".



Graph 5. – Types of documents retrieved in all GLs

In Graph 5 we can observe that a significant percentage of entries of the vocabulary is found in all GL lexicons as well: "abstract", "analysis", "annual", "article", "bibliography", "catalogue", "conference paper", "directory", "dissertation", "index", "interview", "journal", "map", "monograph", "proposal", "protocol", "report", "review", "software", "standard", "study", "thesis", "website".
At the end of this terminological overview based on the *Vocabulary of the types of Grey Literature* these are the entries of the dictionary which cannot be found in our corpus: "bachelor's thesis"; "call for papers"; "codebook; "conference materials"; "conference proceedings"; "course text"; "exam topics"; "green paper"; "house journal"; "master's thesis"; "minutes"; "product catalogue".

Conclusions

To conclude, this survey on the results of the information extraction process performed by the described NLP tool has been a sort of linguistic path in the past and present of the terminology used in GL proceedings with the goal of drawing a picture of the lexicon used by the GL community and thus contributing to get a deeper knowledge of the GL domain.
Many of the terms encountered cannot have synonyms because they reflect specific concepts devoid of the ambiguities peculiar to the common language. Some expressions

such as “grey resources” and “open access” or nouns as “library” and “repository” refer straight and univocally to the “documentary science”, that is they belong to a specific semantic field.

By adopting a diachronic point of view, a significant terminological stability can be noticed. However some terms have been pointed out as obsolete while others emerged as very up-to-date, the latter are those chosen for assembling studies in the same domain or even for labeling emerging fields of knowledge. This is the case, for example, of the bigram “electronic dataset” retrieved in 2004 and 2007 glossaries and then substituted by the bigram “digital dataset” in 2010 and 2014.

Examples could be endless but the size of the corpus had made necessary to delimit the study to a sample by choosing some of its parts and pertaining taxonomies.

In these last twelve years we have witnessed the establishment of new paradigms of scientific communication, the stunning development of information technology and the creation of new infrastructures for storing, preserving and disseminating scientific information. A fact clearly comes to light from this analysis: the grey literature field has a dynamic and cross nature, its community is sensible to technological innovation and proves to be able of keeping pace with the changes.

The lexicon adopted in the GLs’ scientific papers has confirmed that the “grey” community soon paid specific attention to topics like “open access”, “repository” , “digital objects” and “preservation”, just to cite a few. At the same time the almost stable use of a technical and specialized terminology over the time indicates the interest and the willingness to deepen the knowledge of some themes by reporting updates and novelties.

Lastly, this work must be considered a preliminary analysis of the GL corpus, a linguistic resources to be further investigated with different purposes and different tools.

Essential References

1. Bourigault D., Jacquemin C., L’Homme Marie C. (2001). Introduction. In *Recent Advances in Computational Terminology*, VIII-XVIII. Amsterdam, John Benjamins.
2. Cabré M.T. (2000). La terminologia tra lessicologia e documentazione: aspetti storici e importanza sociale, http://web.tiscali.it/assiterm91/cabreita.htm.
3. Cassese S. (1992). Introduzione allo studio della formazione. In «Rivista trimestrale di diritto pubblico» 2, 307-330.
4. Goggi S., Pardelli G., Sassi M., Giannini S., Biagioni S. (2015). A Terminological Survey on the Titles of the Seventh Framework Programme (FP7). In *Proceedings of the Fourteenth International Symposium on Comunicación Social: retos y perspectivas*, vol. I, 223-227. Centro de Lingüística Aplicada, Ministero de Ciencia, Tecnología y Medio Ambiente, Cuba.
5. Goggi S., Monachini M., Frontini F., Bartolini R., Pardelli G., De Mattei M., Bustaffa F., Manzella G. (2015). Marine Planning and Service Platform (MAPS): An Advanced Research Engine for Grey Literature in Marine Science. In *Proceedings of the Sixteenth International Conference on Grey Literature (GL16)*, 108-115. TextRelease, Amsterdam.
6. Goggi S., Pardelli G., Bartolini R., Frontini F., Monachini M., Manzella G., De Mattei M., Bustaffa F. (2016). A semantic engine for grey literature retrieval in the oceanography domain. In *Proceedings of the Seventeenth International Conference on Grey Literature (GL17)*, 104-111. TextRelease, Amsterdam.
7. Kageura K. (1998/1999). Theories of terminology: a quest for a framework for the study of term formation. *Terminology* 5 (1) 21-40.
8. Megerdoomian K. (2003). Text Mining, Corpus building, and testing. In *Handbook for Language Engineers,* 213-268. CSLI Publications, Stanford.
9. Pardelli G., Goggi S., Giannini S., Biagioni S. (2016). Two decades of terminology: European framework programmes titles. In *Proceedings of Tenth International Conference on Language Resources and Evaluation. (LREC 2016)*, 373-378. ELRA - European Language Resources Association, 2016.
10. Park Y., Byrd Roy J., Boguraev Branimir K. (2002). Automatic Glossary Extraction: Beyond Terminology Identification. In *Proceedings of the 19th international conference on Computational linguistics (COLING '02)*, http://aclweb.org/anthology/C02-1142.
11. Pazienza M.T, Virdigni M. (2003). Agent based ontological mediation in IE systems. In *Information Extraction in the WEB Era. Natural Language Communication for Knowledge Acquisition and Intelligent Information Agents*, 92-128. Springer, Heidelberg.
12. Pejšová P., Vaska M. (2010). An Analysis of Current Grey Literature Document Typology. In *Book of Abstracts of the 12th International Conference on Grey Literature (GL12)*, 39-47. Text Release, Amsterdam.
13. Pejšová P., Simandlová T., Mynarz J. (2012). A linked data Vocabulary of the Types of Grey Literature: Version 1.0. In *Proceedings of the Thirteen International Conference on Grey Literature (GL13)*, 170-173. TextRelease, Amsterdam.

Appendix 1 – Frequency

High segment													
Term	**GL5**	**GL6**	**GL7**	**GL8**	**GL9**	**GL10**	**GL11**	**GL12**	**GL13**	**GL14**	**GL15**	**GL16**	**Total**
Literature	405	604	277	252	527	263	160	466	403	363	143	254	4117
Information	433	344	455	264	456	317	298	210		355	497	277	3906
Grey	421	579	275	267	520	299	196	515		366	146	267	3851
Research	294	266	314	153	269	250	193	192	403	532	508	223	3597
Document	260	360	392	118	332	143	201	168		155	168	115	2412
Library		299	276	152	188	312	123	267		153	73	91	1934
Access		152	310	130	136	137	133	112		148	231	198	1687
Report	315	193	165	94			161	197				184	1309
Datum								144		358	367	257	1126
System		158	186		156			117			227	76	920
Publication	230	131		107	233						213		914
Repository		157	187		129	181						142	796
Project		183	164	168							271		786
Open			144	80		159					190	153	726
Collection		213	152	96			102	155					718
Journal		139			176			98			153		566
Science					129					141	201	84	555
Digital		188	180					110					478
Material		146				126	109						381
Metadata		147	137	92									376
User		140						114				73	327
Thesis			141			152							293
Citation		153			134								287
Policy		121										116	237
Database		121		102									223
Source		179											179
Technology											158		158
Service			153										153
Development			130										130
Indexing								122					122
Resource				122									122
Quality								98					98
License												91	91

Table 1

Medium segment													
Term	GL5	GL6	GL7	GL8	GL9	GL10	GL11	GL12	GL13	GL14	GL15	GL16	Total
Datum	80	119	125	65	106	106	88		229				918
Project	130				121	76	95	64	139	129		67	821
User	72		90		104	69	86		136	104	122		783
Repository	48			70			86	85	299	94	84		766
Service	54	69			65	69		84	106	126	125	63	761
Development	93	95		62	61	70	47	63	87	79	101		758
Digital	75			60	66	80	44		166	99	106		696
Collection	97				48	109			198	75	67	69	663
System	130			68		112	96		146	108			660
Science	141	85	64	46		63	53	96	107				655
Resource	36	83	130			60	60		87	112	82		650
Technology	91	73	64			66	45		124	113		61	637
Web	124	84		51	51	97	43		55	87			592
Database	92		90		86	91	64		51	50	65		589
Social	81				85				254	62	82		564
Report					95	106			116	117	128		562
Material	107		82	66	95				56	77	75		558
Process	57	63	110					60	57	107	88		542
Source	39	80	68		65		60		100	69		55	536
Knowledge	62		51	51			39		87	107	138		535
Open	51	78			70		74	67	92	88			520
Community	38		68		78		40		97	109	85		515
Management	52			64				67	54	87	104	69	497
Publication			100			94	48	60	66	120			488
Archive		67	94			116	56		93	43			469
Article		86	87	53	97					52	88		463
Library	158								221		82		461
Format	55	82	68			47			62	44	99		457
Electronic	66	65	85		60	68				44	66		454
Metadata	46				51	88	91		95		79		450
Journal	92		92	67						115		61	427
Institutional	53	69	90			48			101			57	418
Grey									394				394
Survey	71			53	53	69			50	74			370
Academic	72	72	58		60	48						55	365
Communication	65				94				108			54	321
Policy	69		73	68		48				53			311
Online	50	51	51						56	42	54		304
Standard	46	60	68							57	58		289
Citation	143		54	63									260
Access	101								140				241
Health		63	110	58									231
Education	55					65				107			227
Dissertation	45				48	107							200
Model	59		74								63		196
Environment	71								54	70			195
Network	53								81		55		189
Thesis	39	65			85								189
Product	56									129			185
Government		64				56	63						183
Production	90					48				43			181
Website	43									57	81		181
Life					121					58			179
Quality			72						53		54		179
Document									176				176
Book	49		93										142
Documentation	140												140
Data										65		64	129
Practice										47	77		124
Copyright						122							122
Bibliographic	60						53						113
Innovation										113			113

Term	GL5	GL6	GL7	GL8	GL9	GL10	GL11	GL12	GL13	GL14	GL15	GL16	Total
Right							49					62	111
Internet	62									45			107
History				104									104
Security			54	46									100
Discipline					49					46			95
Tool	38			57									95
Review								93					93
Risk									83				83
Patent											82		82
Dataset									80				80
Blog										79			79
Guideline											74		74
Evaluation			73										73
Reactor		66											66
Commercial	62												62
Environmental	62												62
Networking									60				60
Multimedia						59							59
European	58												58
Law							58						58
Application											57		57
Structure			56										56
Corpus										55			55
Training	49												49
Workflow										48			48
Traditional						47							47
Conventional					45								45
Several	44												44
Dissemination	42												42
Engineering	41												41
Magazine	41												41
Protection	41												41
World	41												41

Table 2

Appendix 2 – GL Conference topics

GL	Conference topics
GL5	Models for Academic Grey, Part I: Specific Approaches
GL5	Research is Grey Dependent
GL5	The Economy of Grey
GL5	Strategies for Academic Grey, Part II: General Approaches
GL5	Search Engines are Growing Grey
GL5	Roadmap of Grey Literature Systems and Services
GL5	Alternative Issues in Grey Literature
GL5	Product and Service Reviews
GL6	Institutional Repositories
GL6	Use Analysis
GL6	IT & Research Initiative
GL6	Knowledge Management and Dissemination
GL6	Collection Development and Resource Discovery
GL7	Curriculum Development and Research On Grey Literature
GL7	Theses and Dissertations
GL7	Repositories and Collections of Grey Literature
GL7	Quality Assessment of Grey Literature
GL8	Collection Development, Collection Policies, and Collection Rescue
GL8	Metadata Schemes, Repositories, Software, and Standards
GL8	Curriculum Development and Grey Literature
GL8	Metadata Schemes and Repositories for GL
GL8	Quality Assessment of Grey Literature
GL8	Economic and Legal Aspects of Grey
GL8	Mapping Grey Resources for Costal and Aquatic Environments
GL9	Grey Foundations in Information Landscape

GL	Conference topics
GL9	Tools for Publishing, Archiving, and Accessing GL
GL9	Use and Impact of GL in Scholarly Communication
GL9	Information Walk-Thru, Poster Presentations & Product and Service Reviews
GL9	Grey Literature in Central and Eastern Europe
GL9	New Discoveries in GL for Research Communities'
GL9	Education and Grey Literature
GL9	Information Walk-Thru Poster Presentations, P&S Review
GL10	Institutional Repositories and Grey Literature
GL10	Grey Literature in Biomedical Communities
GL10	Legal Aspects, Intelligence, and Text Mining In Grey Literature
GL10	Grey Literature in Research
GL11	Impact of Grey Literature on Net Citizens
GL11	Uses and Applications of Subject Based Grey Literature
GL11	Grey Literature Repositories
GL11	Open Access to Grey Resources
GL12	Redefining Grey Literature
GL12	New Stakeholders in Grey Literature
GL12	Standardization in Grey Literature
GL12	New Frontiers in Grey Literature
GL13	Social Networking
GL13	Special Collections
GL13	Open Access and Wealth Creation
GL13	Data Frontiers
GL14	Tracing the Research Life Cicle
GL14	Tracking Methods for Grey Literature
GL14	Adapting new Technologies
GL14	Repurposing Grey Literature
GL15	Technology Assessment
GL15	Sustaining Good Practices
GL15	Research and Data
GL15	Towards Informed Policies
GL16	Public Awareness of Grey Literature
GL16	Publishing and Licensing Grey Literature
GL16	Open Access to Research Data
GL16	Managing Change in Grey Literature

Table 3

Appendix 3 - Types of documents

Vocabulary terms	GL5	GL6	GL7	GL8	GL9	GL10	GL11	GL12	GL13	GL14	GL15	GL16	Total
Report	315	193	165	94	95	106	161	197	116	117	128	184	1871
Journal	92	139	92	67	176	44	24	98	35	115	153	61	1096
Study	111	96	67	23	102	74	33	66	83	75	110	77	917
Thesis	39	65	141	9	85	152	12	31	33	14	51	25	657
Article	29	86	87	53	97	24	30	42	26	52	88	32	646
Analysis	57	65	33	27	46	38	45	75	58	77	47	48	616
Standard	46	60	68	19	37	24	21	33	48	57	58	16	487
Website	43	29	20	14	31	38	30	46	41	57	81	32	462
Dissertation	45	21	36	4	48	107	7	25	35	12	25	17	382
Review	21	47	32	18	17	10	4	93	35	30	19	15	341
Software	15	30	48	22	15	35	13	21	28	25	49	6	307
Dataset		10	3	2	20	11	4	25	80	2	7	32	196
Annual	21	2	7	6	8	16	19	76	11	8	5	3	182
Abstract	18	24	24	6	13	11	22	9	12	18	9	10	176
Questionnaire	1	15		19	16	34	13	8	15	24	19	1	165
Index	32	32	16	11	16	8	4	7	12	8	6	11	163
Patent	16	3	9	2	6	5		3	7	14	82	11	158
Catalogue	21	18	20	5	23	22	5	5	3	2	9	4	137
Bibliography	4	15	8	2	10	30	28	1	1	6	18	3	126
Annual Report	7		6	3	3	12	4	66	1	4	1		107
Protocol	13	28	15	3	4	6	12	3	3	5	12	2	106
Proposal	26	16	10	5	7	7	1	6	5	6	5	7	101
Interview	9	16	6	5	5	4	12	2	4	14	15	3	95

Vocabulary terms	GL5	GL6	GL7	GL8	GL9	GL10	GL11	GL12	GL13	GL14	GL15	GL16	Total
Monograph	23	3	9	2	9	3	2	2	3	25	8	4	93
Map	7	8	2	7	2	8	12	3	9	5	13	6	82
Conference Paper	4	17	9	1	7	7	12	5	3	4	4	3	76
Preprint	12	10	10		9	6		4	16		2	4	73
Directory	4	6	10	7	6	7	2	1	12	7	1	7	70
Newsletter	22	15	4	5	9	3			5	4	1		68
Manual	5	4	1	5		9	7	15	2	6	2	1	57
Bulletin	13	3	2	3	3	4			1	1	2	2	34
Curriculum	3		4	2	6	2		7	2	4		1	31
Poster	1		1	2	1	1	5	7	2	1	5	4	30
Brochure	11		2	2	4	1	1	2			1		24
Proceedings	1	6	1		1	1		1	6	4	1		22
Government Document	1	5		2	4	1	3	2				1	19
Glossary	1	1	1			1	1		2		6	3	16
Memorandum			2			1	5	1	1	1	1		12
Handbook	2		2	3	1			1		1	1		11
Timeline			3	2					2	1		3	11
Announcement	1				3			1	2		2	1	10
Conference Program					1			1	5		1	1	9
Essay	2				1	2	2		1	1			9
Press Release			1			5					1	1	8
Chronicle			2						2	1		1	6
Leaflet					3	1		1				1	6
Course Material	2						1	1	1				5
Informative Material									1		3	1	5
Normative Document					2	1					2		5
Anthology						1	1		1	1			4
Research Plan									1	1	2		4
Syllabus					3								3
Tertiary Source	1				1			1					3
Corporate Literature											2		2
Habilitation Thesis									1		1		2
Image Material											2		2
Legal Document							1				1		2
Guidebook		1											1
Technical Documentation		1											1

Table 4

Altmetrics and Grey Literature: Perspectives and Challenges

Joachim Schöpfel, GERiiCO Laboratory, University of Lille
Hélène Prost, CNRS France

Abstract
Traditional metrics largely overlook grey literature. The new altmetrics introduced in 2010 as "new, online scholarly tools (that allow) to make new filters" (Altmetrics Manifesto), can include all kinds of scholarly output which makes them interesting for grey literature. The topic of our paper is the connection between altmetrics and grey literature. Do altmetrics offer new opportunities for the development and impact of grey literature? In particular, the paper explores how altmetrics could add value to grey literature, in particular how reference managers, repositories, academic search engines and social networks can produce altmetrics of dissertations, reports, conference papers etc. We explore, too, how new altmetric tools incorporate grey literature as source for impact assessment, and if they do. The discussion analyses the potential but also the limits of the actual application of altmetrics to grey literatures and highlights the importance of unique identifiers, above all the DOI. For the moment, grey literature missed the opportunity to get on board of the new movement. However, getting grey literature into the heart of the coming mainstream adoption of altmetrics is not only essential for the future of grey literature in open science but also for academic and institutional control of research output and societal impact. This can be a special mission for academic librarians.

Introduction
Traditional metrics largely overlook grey literature. Worse, they basically disregard grey literature as irrelevant for the evaluation of research. Established metrics for individuals and organisations are journal-centric. Measuring the performance and popularity of scientists or research structures means counting the number of articles citing other articles, resulting in journal impact factors, normalized citation rates and the h-index. Even those rare studies including conference papers are limited to published proceedings[1]. Grey literature remains out of scope. The most important reason is the way these metrics are produced – they rely on bibliographic tools like the Web of Sciences (WoS) and Scopus which from the beginning on were (nearly) exclusively journal and monograph A&I services, dismissing other vectors of scientific communication outside of the academic publishing market[2].

The emergence of webometrics, i.e. the "study of the quantitative aspects of the construction and use of information resources, structures and technologies on the web drawing on bibliometric and informetric approaches" (Björneborn & Ingwersen 2004, p. 1217), change the situation. As many scholarly activities today are web-based, the field of webometrics is partially covered by scientometrics (figure 1). These new or alternative metrics are not limited to journals but apply to academic content (scholarly work) at large, insofar and as long as this content is available on the web, in particular on the social web (Galligan & Dyas-Correia 2013). They are sometimes called scholarly metrics or social media metrics, and most often defined as altmetrics.

The fact that these new metrics can include all kinds of scholarly output makes them interesting for grey literature. In a draft on altmetrics definitions and use cases, the National Information Standards Organization describes scholarly output as "a product created or executed by scholars and investigators in the course of their academic and/or research efforts. Scholarly output may include but is not limited to journal articles, conference proceedings, books and book chapters, reports, theses and dissertations, edited volumes, working papers, scholarly editions, oral presentations, performances, artifacts, exhibitions, online events, software and multimedia, composition, designs, online publications, and other forms of intellectual property" (NISO 2016, p.9). One part of this output clearly belongs to grey literature, especially when citable and accessible[3].

[1] See for instance Ingwersen et al. 2014, also for similar, older studies
[2] The methodological problems to identify theses in bibliographic databases in Larivière et al. (2008) confirm the situation
[3] See the definition of "acceptable products" by the National Science Foundation, *Grant Proposal Guide II-12 NSF 14-1*, November 2013 http://www.nsf.gov/pubs/policydocs/pappguide/nsf14001/gpgprint.pdf

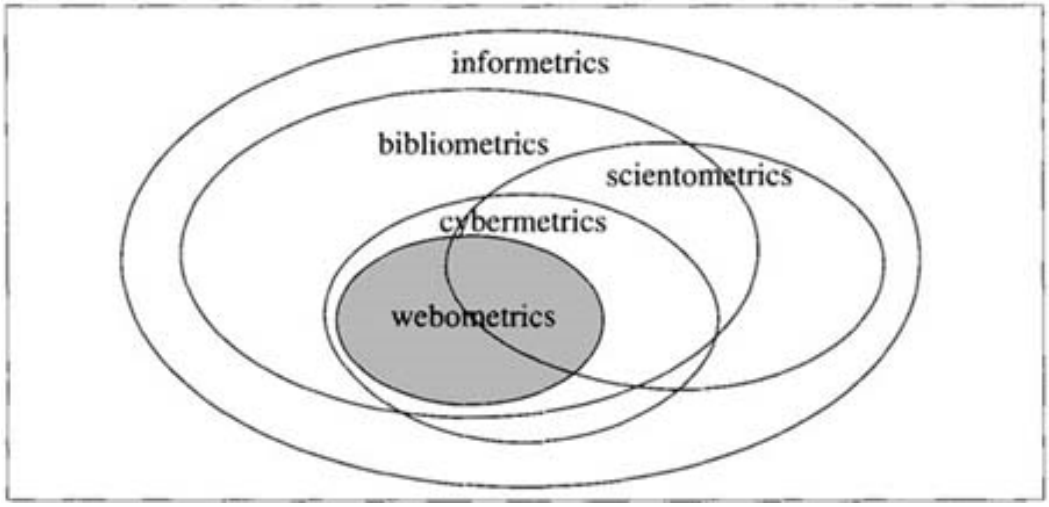


Figure 1: Webometrics in the field of library and information sciences (source: Björneborn & Ingwersen 2004)

The topic of our paper is the connection between altmetrics and conference proceedings, reports, theses and dissertations, and working papers. Do altmetrics offer new opportunities for the development and impact of grey literature? Are there already examples of good practice? Are there any barriers? However, before we outline the potential of altmetrics for grey literature, we will provide some elements for a better understanding of this concept.

A short history of altmetrics

Altmetrics have a short history[4]. The term was introduced by Jason Priem from Chapel Hill in 2010, in a tweet published on the 29th September 2010: "I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics"[5]. This came after the global success of the web 2.0 tools and media, such as Facebook, Twitter etc., and it became popular as a kind of marketing umbrella for a broad range of new metrics of scholarly impact on the social web (Priem & Hemminger 2010).

The Altmetrics Manifesto[6] from 26 October 2010 merges article-level metrics and distributed scientific evaluation with social media into research on altmetrics and defines them as fast and open filters to relevant and significant scholarly sources, not in continuity but in disruption with webometrics or citations; "given the crisis facing existing filters and the rapid evolution of scholarly communication, the speed, richness, and breadth of altmetrics make them worth investing in" (Priem et al. 2010).

From that moment on, the interest for altmetrics increased steadily to join and finally exceed scientometrics, according to Google Trends (figure 2). Two years after the Manifesto, the San Francisco Declaration on Research Assessment (DORA), initiated by the American Society for Cell Biology (ASCB), recognizes the need to improve the ways in which the outputs of scientific research are evaluated and suggests the "use of a range of article metrics and indicators on personal/supporting statements, as evidence of the impact of individual published articles and other research" (DORA 2012). Signed by nearly 12,500 individuals and 800+ organizations[7], DORA fostered the awareness for altmetrics and became a reference for the debate, research and development in the field.

[4] See comprehensive reviews by Erdt et al. (2016) and Sugimoto et al. (2016)
[5] https://twitter.com/jasonpriem/status/25844968813 by @jasonpriem
[6] http://altmetrics.org/manifesto/
[7] http://www.ascb.org/dora/ accessed 7 September 2016

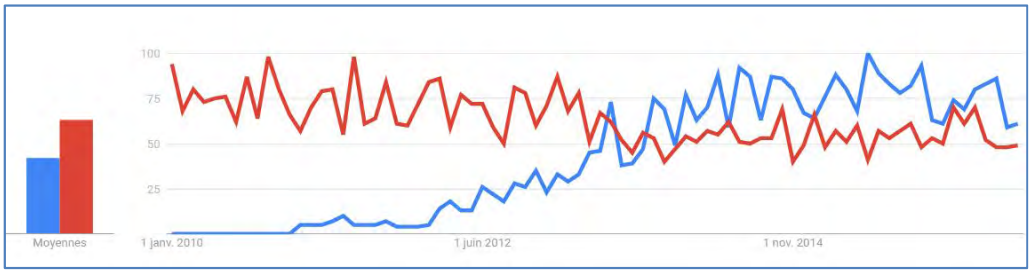


Figure 2: Altmetrics (blue) on Google Trends, compared to scientometrics (red) (2010-2016)[8]

The increasing number of scholarly work dedicated to altmetrics reveals the same trend (figure 3). No study on altmetrics before 2010, and then a steadily growth from 8 references in 2010 to 122 in 2015.

The Google Scholar statistics confirm the Google Trend figures – the interest for scientometrics remains relatively stable, with 50-70 publications per year, but is exceeded by works on altmetrics from 2013 on.

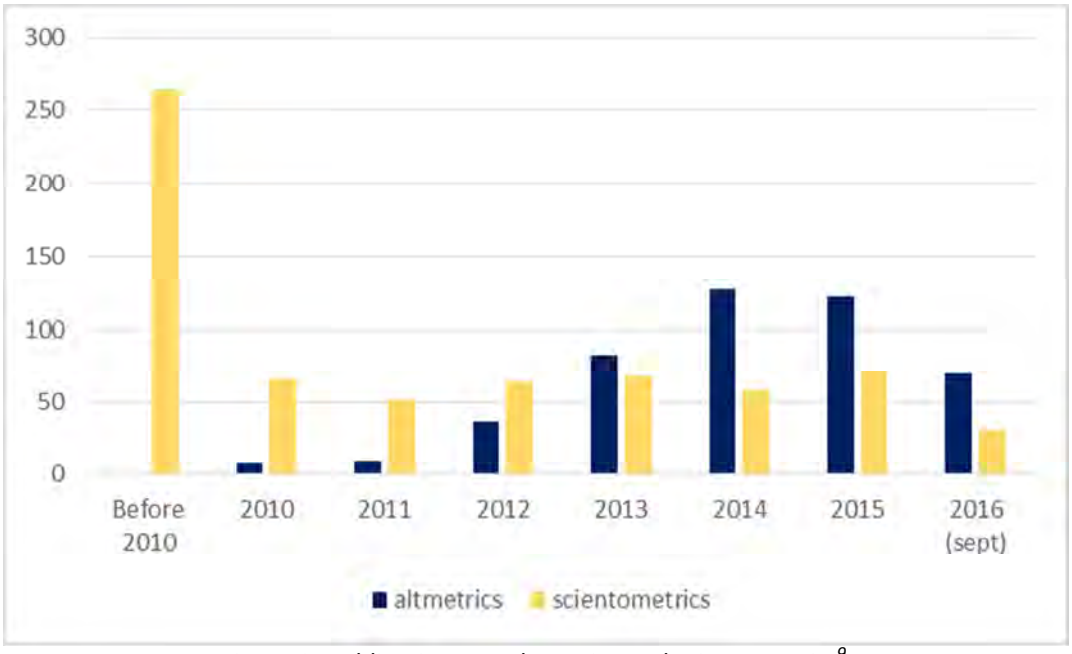


Figure 3: Publications on altmetrics and scientometrics[9]

Basically, altmetrics are "social web metrics for academic publications" (Sud & Thelwall 2014, p.1131) and particularly interesting for measuring societal impact, beyond the academic community (Piwowar 2013), through the count of views, downloads, clicks, likes, tags, posts (blogging) and tweets (micro-blogging), shares, discussions etc. The term "usually describes metrics that are alternative to the established citation counts and usage stats— and/or metrics about alternative research outputs, as opposed to journal articles" (NISO 2014, p.4).

Variety is one main feature of altmetrics, a class of indicators measuring attention, dissemination and influence[10], even if the distinction between attention, dissemination and influence is not self-evident. The main areas of altmetrics are shown in figure 4. Impact on the (social) web can be assessed through the count of PDF or HTML downloads (viewed), the creation of references in online reference managers like CiteULike, Zotero or Mendeley (saved), the number of posts in blogs and micro-blogs, on Facebook or Wikipedia (discussed), the number of mentions in editorials or tools like F1000 (recommended) or as usual, simply via the number of citations in the WoS, Scopus, PubMed Central or CrossRef (cited).

[8] Data source: Google Trends www.google.com/trends accessed Sept 3, 2016
[9] Data source: Google Scholar https://scholar.google.fr allintitle: altmetrics (or scientometrics), accessed Sept 5, 2016
[10] See https://www.altmetric.com/about-altmetrics/what-are-altmetrics/

The NISO Alternative Assessment Metrics Initiative (2016) defines altmetrics as a broad concept that includes "multiple forms of assessment that are derived from activity and engagement among diverse stakeholders and scholarly outputs in the research ecosystem".

Today, a clear, common, widely accepted definition is not in sight. Altmetrics comprise many different types of metrics in a constantly changing landscape and "refer to a heterogeneous subset of scholarly metrics and are a proper subset of informetrics, scientometrics and webometrics" (Haustein 2016, p.416). Perhaps a pragmatic approach like Altmetric's recent definition will fit best, for the moment: "Altmetrics are attention data from the social web that can help librarians understand which articles, journals, books, datasets, or other scholarly outputs are being discussed, shared, recommended, saved, or otherwise used online. They can be reported at the item-, journal-, or author-level"[11].

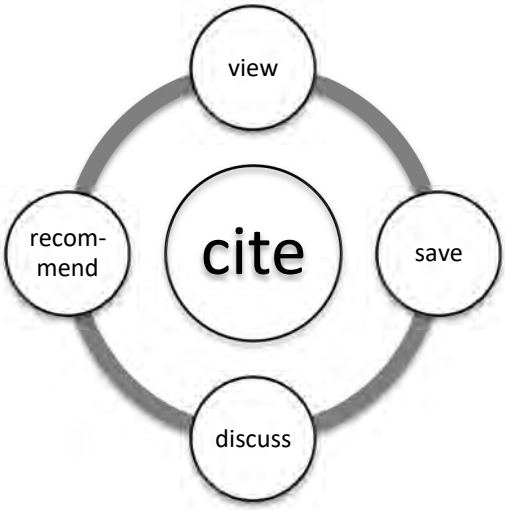


Figure 4: Altmetrics areas of assessment

Six years after the Manifesto, however, it is not quite clear if altmetrics are "an alternative or enhancement to the use of journal impact factors and click-through rate analysis to measure the impact and value of scholarly work" (Galligan & Dyas-Correia 2013, p.56). But they are already relevant for research evaluation. The European Commission DG Research and Innovation has established an Expert Group on Altmetrics which describes the emergence of altmetrics as part of the "transition to a more accountable and transparent research system"[12], more efficient, open to society, and expects "robust, responsible, transparent and interoperable uses of metrics and altmetrics in open science". Altmetrics are levers in support of open science. Up to now, including altmetrics in decisions on grants, hiring and tenure still requires careful consideration but they may soon become a normal part of a CV (Kwok 2013).

What does this mean for grey literature? What is the potential of altmetrics for grey literature? The next section tries to provide a global answer.

The potential

Bornmann (2014) mentions four benefits of altmetrics compared to traditional metrics: they measure impact beyond science, they can include scholarly products other than papers (articles), they allow impact to be measured shortly after the output, and as a rule, it is easy to obtain altmetric data (figure 5).

Compared to traditional, citation-based metrics, altmetrics endorse two different developments: "Widening the definition of research outputs to include more than just books and journal articles, and looking beyond citations for a quantitative way of assessing or discovering them" (Adie 2016, p.67). Thus, at least in theory, altmetrics are not limited to a

[11] https://www.altmetric.com/blog/altmetrics-collection-development/

[12] Next-generation altmetrics: responsible metrics and evaluation for open science, available at https://ec.europa.eu/research/openscience/index.cfm?pg=altmetrics_eg

coverage similar to the WoS or Scopus. As stated by Andy Tattersall, "altmetrics focuses on research artefact level metrics that are not exclusive to traditional papers but also extend to book chapters, posters and data sets among other items" (2016, p.1). "Among other items" – this could or should bring in non-traditional, non-commercial items, like working papers, dissertations, conference papers, reports etc.

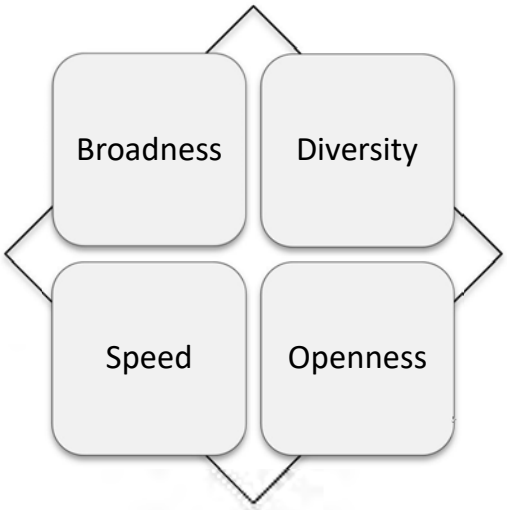


Figure 5: Benefits of altmetrics (source: Bornmann 2014)

And even if altmetrics still focus primarily on practices relating to research articles as "central research output that informs research assessment (they) can and should be extended by recognizing additional products, such as datasets (…)" (DORA 2012). Therefore, their potential for grey literature is twofold (figure 6):

- Diversity: Impact assessment on article level such as download counts also applies to grey literature. "Altmetrics (…) allow for evaluation of a greater diversity of products, i.e., not just publications (…). These products might be datasets, software, copyrights, algorithms, grey literature, and slides (...). Altmetrics now offer the opportunity to determine the impact of these products both in science (…) and beyond science" (Bornmann 2014, p.898). Diversity, as said above, is considered as one crucial advantage of altmetrics, and this includes grey literature.

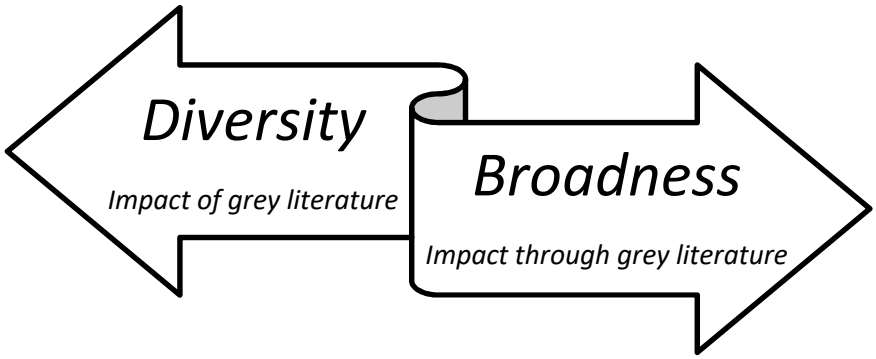


Figure 6: Double potential of altmetrics for grey literature

- Broadness: In contrast to traditional metrics which usually exclude "documents such as technical reports or professional papers which some label as 'grey literature' (…) due to lack of indexing" (Moed & Halevi 2014), altmetrics are not limited to scientometric databases; also their impact assessment based on citations, links and referrals can take account of a broader range of scientific information, including citations in dissertations, reports, white papers etc. Thus, grey literature can serve as material to measure impact of scientific output.

Beyond these two major perspectives for grey literature, recent studies on altmetrics mention two other potential benefits:

1. Dissemination of reports: Altmetrics may be a way to foster more efficient information practices by research organizations and funding bodies (foundations); one obstacle to disseminate reports etc. via open and shared systems is that often "organizations aren't sure people are even reading this stuff (…) Altmetrics (…) hopefully can better inform our expectations and measures of readership" (Brooks & Fitz 2015, p.43; see also Dinsmore et al. 2014).
2. Scientific information in developing countries: Neylon et al. (2014) insist on the application of altmetrics, especially social media usage metrics, for grey literature in "a developing country context, such as in sub-Saharan Africa (where) the importance of 'grey literature' – policy briefs, working papers, media articles and other scholarship aimed at lay audiences – is massive, satisfying both the need for social engagement as well as scholars' professional expectations" (p.2).

In the following, we will address the first two issues, diversity ("altmetrics for grey literature") and broadness ("altmetrics through grey literature"), with some examples and a focus on special conditions and prerequisites.

Altmetrics for grey literature

Our first issue is about impact assessment of grey literature. As said above, traditional metrics have largely overlooked grey literature. Altmetrics can offer new and unique opportunities for the web-based impact measurement of reports, conferences, dissertations etc. But do they really? And if so, how? To what extent?

In 2012, a study funded by the Dutch SURF-Foundation assessed fifteen "novel impact monitors", such as reference managers, academic search engines and new altmetrics tools (Wouters & Costas 2012). At least nine out of the fifteen "monitors" can produce impact data for grey literature[13]. These seem rather favourable and promising conditions. Let's get some empirical insight for a better understanding.

Repositories

In the GreyNet community, repositories, especially institutional repositories, are generally considered as "natural home" for grey literature, as institution-based platforms for the dissemination and preservation of the institutional scientific output (Banks & de Blaaij 2006). Most of the open repositories contain one or more categories of grey literature, often theses and dissertations (particularly in university repositories), but also conference and workshop papers, unpublished reports, working papers or other "special items"[14].

All repository servers produce log files of views and downloads which can be transformed into statistics and metrics, useful as well for institutions and hosting organisations as for authors and readers. However, a couple of years ago only few repositories made these metrics freely available on their website, along with the metadata and deposited files, and even less did so in a standard, interoperable way (Schöpfel & Prost 2009, Prost et al. 2010).

The debate on new metrics accelerated the movement, and following the Altmetrics Manifesto and the DORA Declaration, repository hosts and managers started to improve the availability of web analytics and to implement new altmetrics tools. As a result, today "institutional repositories are (…) embracing altmetrics as a means of both tracking and encouraging engagement with the resources, and the ability to track and measure engagement with grey literature can be a good source of evidence of the role these outputs play in the research and publication life-cycle" (Priego 2014). Four very different examples may show the potential but also the limits of this development.

[13] GS, MAS, ArnetMiner, Mendeley, CiteULike, Zotero, ReaderMeter, ImpactStory, SURE2

[14] See the Directory of Open Access Repositories OpenDOAR, available at http://www.opendoar.org/

HAL[15]

The French national repository HAL contains 400,000+ documents and 1,1m records. Nearly half of the full text deposits (46%) are grey literature, with nearly 60,000 dissertations and more than 80,000 conference papers. For all these documents, HAL produces usage statistics on the item level, of full text downloads and retrievals (views) of the records (metadata). Since the launch of HAL in 2001, authors as well as collection managers have access to detailed and customizable usage statistics for each item or, cumulated, for a collection, an institution, an author etc.

On the public interface, HAL displays for each record two metrics, cumulated metadata views and full text downloads. In our example (figure 7), HAL shows that the Lille 3 White Paper on research data in PhD theses received so far 3,591 record views and 2,150 successful download requests[16]. But HAL does not offer comparative metrics (average statistics per document type and/or domain etc.).

Since 2015, HAL displays an Altmetric badge with metrics based on the unique identifiers DOI, arXiv-id and PubMed ID. So far, HAL does not allocate DOIs to deposits without identifiers and does not use its own identifier HAL Id or other identifiers like the French national dissertation number (NNT) for the assessment of altmetrics. Thus, the only conference papers with Altmetric badges we could find in HAL are those published by Springer, IEEE or other commercial publishers specialised in proceedings and members of CrossRef. Probably, this means that while all grey literature in HAL is displayed with usage statistics, no grey item has received an Altmetric badge up to now.

Figure 7: Display of usage statistics in a HAL record

figshare[17]

The online digital repository figshare where researchers can preserve and share their research outputs contains above all figures, datasets and filesets but also some papers, dissertations, posters and presentations. Figshare has "three basic functions: it acts (1) as a personal repository for yet unpublished materials, (2) as a platform for newly published research materials, and (3) as an archive for PLOS" (Kraker et al. 2015). In fact, almost 90% of the input comes from PLOS – mostly figures, while text files represent less than 2% of all entries, and the part of dissertations (all kinds of short or long unpublished written texts), posters and presentations is extremely low (0.3%).

Figshare exhibits view and download counts for all deposits. In April 2016 figshare implemented Altmetric badges to showcase attention surrounding research output (figure 8).

[15] HAL = Hyper articles en ligne https://hal.archives-ouvertes.fr/
[16] Accessed 9 September 2016
[17] https://figshare.com/

Figure 8: USAGE statistics and altmetrics in figshare

The reader can click through to detailed impact information on the Altmetric server. All figshare entries have a DOI; 89% of the DOIs are provided by PLOS, the other 11% are allocated by figshare (with DataCite) as soon as a user makes an uploaded material publicly available. This means that dissertations may get a DataCite DOI[18]. The systematic allocation of DataCite DOIs facilitates the generation of attentions scores with the Altmetric tool. However, obviously not all figshare entries have an Altmetric badge.

IRUS-UK[19]

IRUS-UK is a Jisc-funded national aggregation service which collects raw usage data from 113 institutional repositories and transforms them into COUNTER-compliant statistics. Insofar as these repositories contain unpublished grey items, IRUS-UK aggregates usage statistics from more than 200,000 conference papers, reports, dissertations and working papers which represent 34% of the repositories' content (figure 9). All these deposits received nearly 30m successful download requests, or 43% of all aggregated IRUS-UK downloads. These statistics are interesting for three reasons.

- Often grey literature usage statistics are not standardized. Here, as the IRUS-UK statistics are COUNTER-compliant, the data are comparable, authoritative, and standard-based.
- For this reason, they can be compared to download figures from other document types, in particular with article statistics. This direct comparison reveals for instance that in early September 2016, the average downloads of journal articles are similar to conference items, proceedings or reports, but two times lower than the usage statistics of dissertations and working papers. Taken together, the average usage of grey literature is one third higher than for articles or books.
- The aggregated usage statistics should allow for further standardization, e.g. for document- and/or domain-specific average download figures that could be used as a kind of reference set for individual items, like the PLOS metrics.

Document type	Total number	Total downloads
Conference or Workshop Item - Other	50 892	4 536 836
Conference Papers /Posters	7 823	389 402
Conference Proceedings	5 256	252 382
Report	14 265	1 704 870
Thesis or dissertation	123 014	21 409 166
Working Paper	5 915	795 653

Figure 9: Grey literature in IRUS-UK[20]

[18] See the following example, a two-page dissertation: Harper, Danny (2016): Plagiarism in college essays and assignments.docx. figshare. https://dx.doi.org/10.6084/m9.figshare.3528812.v1 Retrieved: 10 47, Sep 09, 2016 (GMT)
[19] Institutional Repository Usage Statistics UK http://www.irus.mimas.ac.uk/
[20] All IRUS-UK figures and statistics accessed 7 September 2016

CORE[21]

For five years now, the CORE project aggregated and enriched content from nearly 1 000 repositories from all over the world, in order to increase the discoverability and reusability of open access papers (Pontika et al. 2016). As CORE harvests not only the repository metadata but also the full-text and caches this PDF version in its own database, it can provide IRUS usage statistics on full-text downloads. Among the 37m harvested items (called "articles" or "manuscripts"), CORE also contains reports, conference papers and other unpublished documents. However, the CORE portal does not allow for document-type specific browsing or search, a fact which reduces its interest for our purpose. -

To sum up, these four examples confirm the potential of repositories for the production of altmetrics, on a continuum from usage statistics (views, downloads) to impact measures based on social media and the possibility to display standard-based data and reference sets. The limits or pre-requisites are the need for rich metadata, including the document type, and the allocation of an established unique identifier.

Social networks

Much has been said about academic social networks like ResearchGate and Academia, about their functionalities, their uptake by the research communities[22] and their impact on scientific communication[23], their competitive strategy challenging above all institutional repositories, and their business model. Because of the increasing number of users, records, documents and other material shared via these networks, they are part of these "novel impact monitors" mentioned above. At least three different aspects are relevant for altmetrics: the creation of metadata, the deposit of the document and the mention of a document in a debate or in an answer to a question.

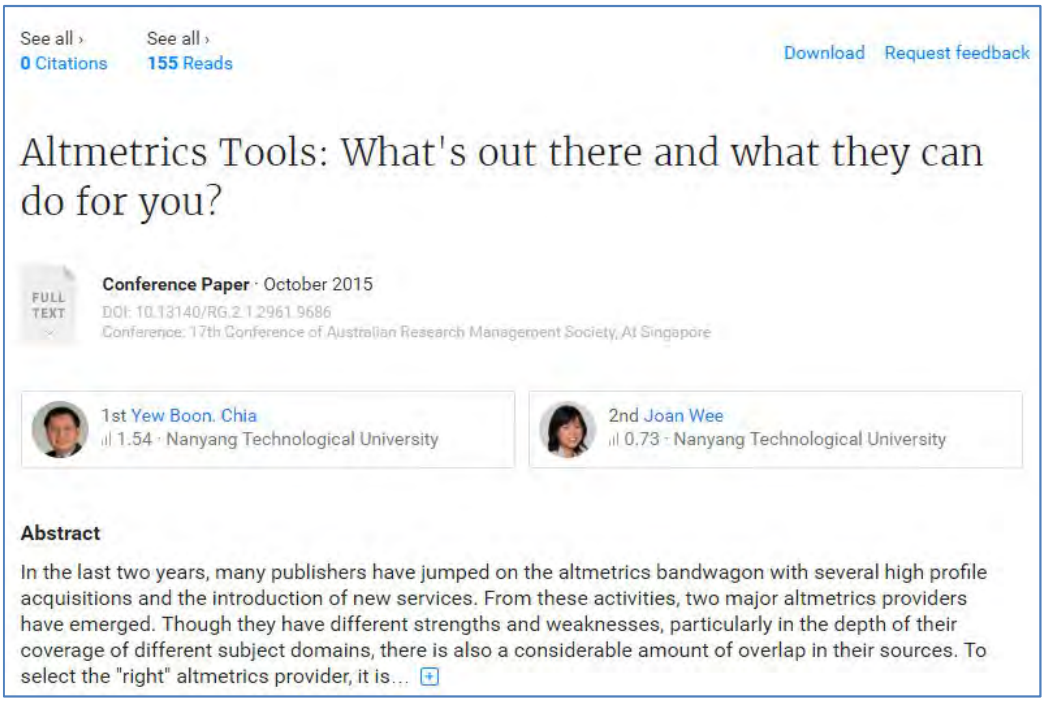


Figure 10: A shared conference paper with DOI (ResearchGate)

Basically, academic social networks invite researchers to share their results, without imposing limits or specific items, i.e. all major types of grey literature can be deposited in social networks. ResearchGate for instance suggests 18 types of "publications", including conference papers, posters, presentations, technical reports, theses and working papers; but also unpublished articles (preprints) and working copies, datasets, negative results and raw data. Also, if necessary, a new format (category) can be created for a specific deposit. Clearly

[21] COnnecting REpositories https://core.ac.uk/
[22] 41m accounts in Academia, 10m accounts in ResearchGate
[23] 14.7m papers in Academia, 100m papers and other items in ResearchGate (24% papers with full-text), mostly STM

they have become large reservoirs for all kinds of unpublished, grey literature, with the potential to make them available for impact measurement. However, this potential is conditioned by the quality of the metadata, in particular of an identifier. The social networks do not allocate unique identifiers but invite to add or import existing DOIs to the deposit (Figure *10).*

Recently, a study was conducted on the effectiveness of six ResearchGate metrics on the author level (ResearchGate score, impact points, number of downloads, number of publication views, number of citations, and number of profile views), concluding that in a small sample and a specific field "the ResearchGate score can be an effective indicator for measuring an individual researcher's performance" (Yu et al. 2016, p.1005).

Beyond academic networks, scientists share and discuss results also on Facebook and Google+; yet, on the one hand these networks are not designed for documentary metadata; on the other hand, the coverage of scientific documents still seems low, producing unreliable metrics (Haustein 2015).

Reference managers

Reference managers like CiteULike, Zotero and Mendeley can provide relevant information for altmetrics, in particular about the number of copies of a given reference. Mendeley for instance is a large database of "white papers, conference proceedings, book and journal references, and other kinds of grey literature that is searchable by other Mendeley users (…)" (Tattersall 2016, p.114). Mendeley provides how many users have a copy for each item.

Figure 11: Number of bookmarks in CiteULike

CiteULike is said to contain 8.3m references and proposes 17 item types, including conference papers, technical reports, Master's and PhD theses, unpublished work and "miscellaneous". There is no available reliable data on the actual number of references for each of these categories. Like Mendeley, CiteULike inform about the number of copies (bookmarks) for each reference (figure 11). CiteULike allocates its own identifier and supports DOI and Pubmed ID, for importing, creating (generating) and searching references.

Bookmarks can be used as a complement to citation metrics. Traditional citation-based indicators, in particular the journal impact factor and author mean citation per paper, are correlated with bookmark-based indicators (altmetrics), such as journal mean bookmarks per paper, the percentage of bookmarked articles in the journal and author mean bookmarks per paper; an analysis of data from the WoS and CiteULike reports the correlations slightly higher for journals than for authors (Sotudeh et al. 2015). Zoller et al. (2016) conclude that a bookmarking system's most inherent feature – tagging – is suitable

for identifying topic subsets of publications where usage and future citations exhibit higher correlations. Yet, apparently no data has been published on bookmark-based metrics of grey literature.

Academic search engines

The SURF-study on altmetrics tools mentions the academic search engines Google Scholar and Microsoft Academic Search because they include "cited by" data for some items, whenever they can identify citations (figure 12). This data can be analysed and interpreted as an indicator for impact on the web.

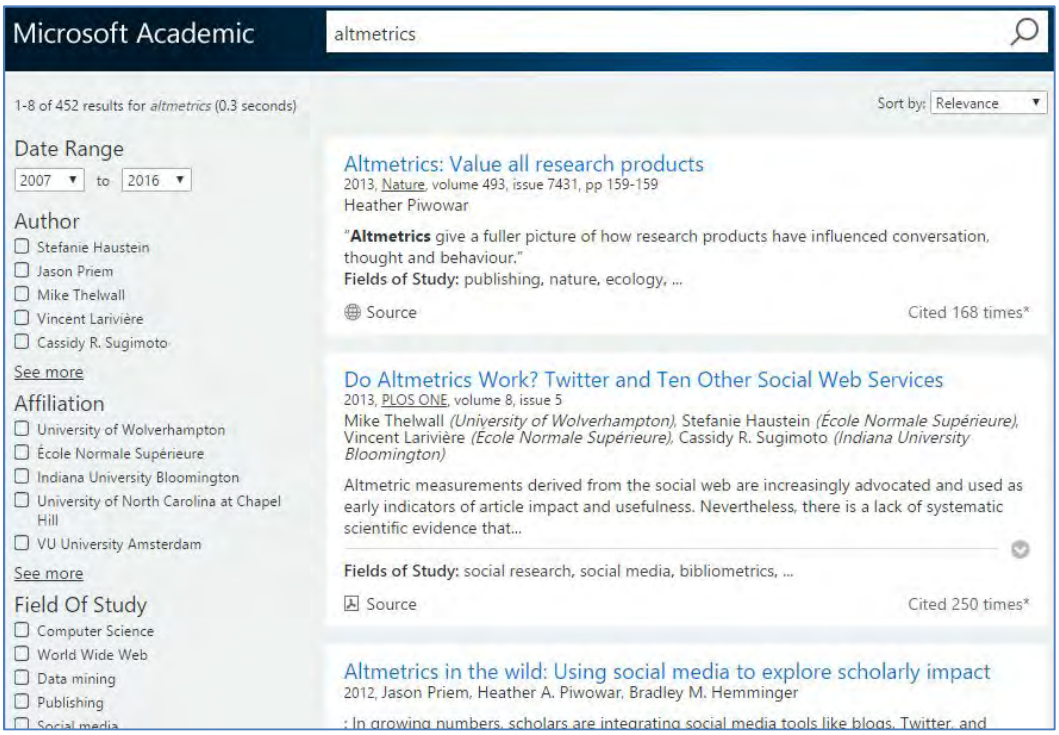


Figure 12: "Cited by" data in Microsoft Academic Search

As these search engines cover a large part of the academic web, in particular institutional repositories and other non-commercial platforms, their crawling and indexing include preprints, dissertations, reports, conference papers etc. For example, figure 13 shows citation data for a workshop paper available on figshare and not published elsewhere, without an allocated DOI.

Obviously, the academic search engine are able to produce impact data for all kinds of scientific papers, as long as they are made available on referenced and indexed platforms, in particular institutional and other repositories. Unique identifiers like the DOI are not indispensable but may improve the reliability of the search results.

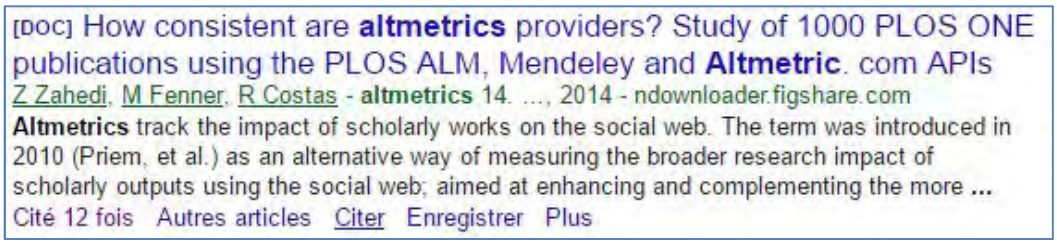


Figure 13: "Cited by" data for a workshop paper on figshare, in Google Scholar

New altmetrics tools

Following Kraker et al. (2015), the most important data provider for altmetrics are not reference managers, academic social networks or search engines but Twitter: "In the altmetrics analysis, we found that Twitter was the social media service where research data gained most attention". A growing number of new tools and platforms aggregate these online events (tweets, likes, comments, downloads etc.) as well as derived metrics from repositories, reference managers etc. (NISO 2016).

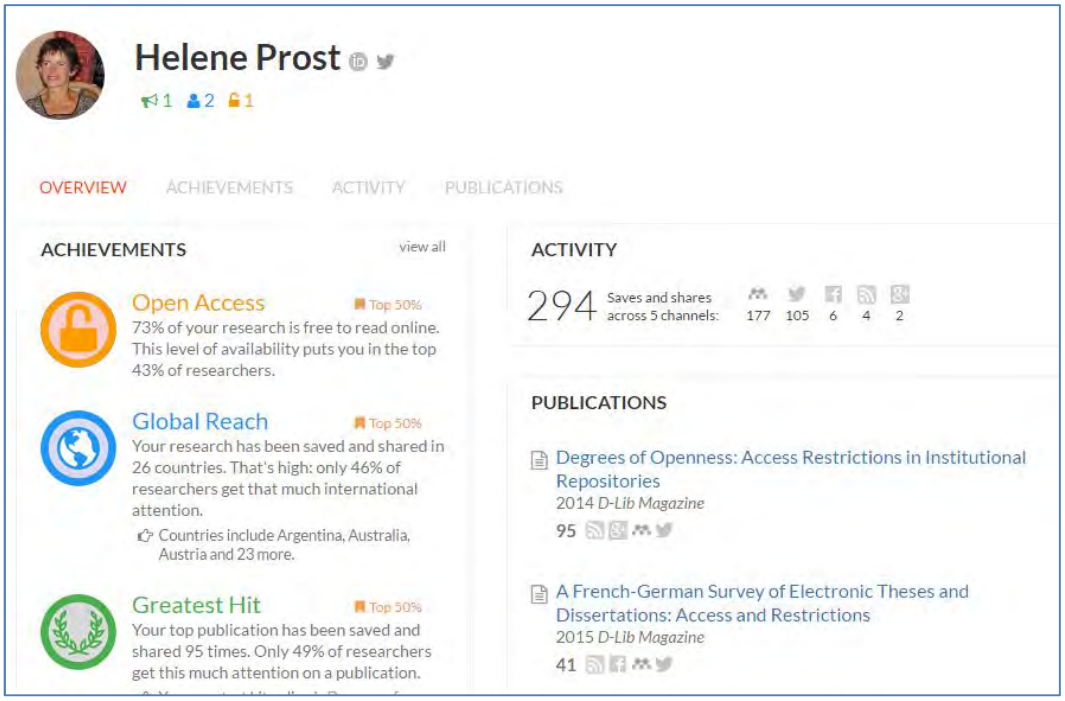


Figure 14: Author's page on ImpactStory

Among these data aggregators one can find Altmetric.com, PlumX, ImpactStory etc. (figure 14). They do not measure the same aspects, and they do not generate the same metrics. While PlumX from PlumAnalytics fits more with libraries' and institutions' needs, especially for repositories (Lindsay 2016), ImpactStory is aimed at individual researchers, and Altmetric offers services for individual researchers, institutions and funders but/and above all for commercial publishers (Konkiel 2012). PlumX detects considerably more items in social media and also finds higher altmetric scores than ImpactStory; but comparison of altmetrics tools is difficult due to differences in assignments to categories, which result in different counts (Kraker et al. 2015).

Basically, these tools can produce social impact metrics for working and conference papers, dissertations and other grey items. However, only few studies have been published on aggregated altmetrics data incorporating grey literature. Altmetric does not track non-traditional outputs. Wilkinson et al. (2014) made use of the Web Impact Report (WIRe) as a novel solution to assess the impact of organisational reports, especially when in open access. WIRe consists of a "range of web-derived statistics about the frequency and geographic location of online mentions of an organisation's reports (…)" (p. 797), such as online citations, site domain and genre of the citing site (blog, governmental sources etc.). Nevertheless, this case study with a small corpus of 20 reports reveals two major issues, i.e. a relatively high percentage of incorrect matches and a time-consuming human workload (content analysis).

ArnetMiner[24] aims to provide comprehensive search and mining services for academic social networks, with a special focus on 6,000+ conferences, mostly in computer sciences, and with a ranking based on the H5-Index, top-cited authors and papers, and data on the social network and semantics for each conference (figure 15).

[24] http://www.aminer.org

Figure 15: Conference analysis of SIAM International Conference on Data Mining on ArnetMiner

Literature about altmetrics mentions other tools like PaperCritic, PeerEvaluation or ReaderMeter. Some are operational, others are not; some are part of PLOS, Mendeley etc. Basically they are all open for grey and other items, primary data etc. But we could not find any reliable information about their real value and interest for grey literature, in terms of specificity and impact.

Altmetrics through grey literature

"Broadness", the possibility to take account of a broader range of scientific information is one of the four major benefits of altmetrics. They could be "more diverse in kinds of data and accordingly numbers of data sources (whereas for traditional citations only the cited references in journals serve as data source)" (Bornmann 2014, p.898), thus revealing more diverse and nuanced forms of impact than traditional indicators. So, how can grey literature contribute to altmetrics? Do altmetrics tools make use of grey literature? For Euan Adie, CEO of Altmetric, grey literature[25] "presents great opportunities for alternative metrics, providing data and indicators that cannot be found anywhere else" but also drew attention to the specific characteristics and challenges, e.g. missing identifiers, no "canonical metadata", lack of long term preservation and availability (Adie 2014). Five examples may illustrate potential benefits and limits of this "broadness".

1. Web-based grey literature can serve as source to increase impact of other grey literature. Wilkinson et al. (2014) conducted their study on WIRe with a small sample of 20 research reports. Their results showed that most of them (17) had been cited by other reports, conference papers, white papers, MA and PhD theses and speeches and/or dissertations available on the web. But without standard or automated procedures, including grey literature involves a lot of human work.
2. In some fields, grey literature may be more relevant than in others. Working on subfields of sustainable energy research, Ingwersen et al. (2014) insist that such analyses "should include proceedings papers – because this document type does have significant (…) influence on the overall citation impact of a research field, in particular in proceedings-dominant fields" (p.1290). The same observation would probably apply to economics (working papers), physics (preprints) and computer science (conference papers).
3. New content mining tools improve the efficiency and broadness of data aggregators. Thus, Altmetric has developed a text-mining solution (*Altmetric Policy Miner*) to discover mentions of publications in policy documents on selected websites[26]. Due to this APM-software, Bornmann et al. (2016) were able to assess the societal impact of climate change publications mainly through grey literature from governmental agencies, international organizations and NGOs.

[25] Described as "theses, posters, preprints, patents and policy documents and similar"

[26] Such as European Food Safety Authority (EFSA), GOV.UK–Policy papers, Research & Analysis, Intergovernmental Panel on Climate Change (IPCC), International Committee of the Red Cross (ICRC), World Health Organization (WHO), International Monetary Fund (IMF), Médicins sans Frontières (MSF), NICE Evidence, Oxfam Policy & Practice, UNESCO and World Bank

4. Human knowledge, manual searching and browsing are the price of "broadness" and inclusion of grey literature in alternative impact assessment. Urquhart & Dunn (2013) evaluated the impact of the National Minimum Dataset for Social Care (NMDS-SC). References to the dataset (citations) were identified in 175 separate publications, with 50% policy and practice reports, 35% media communications and only 15% academic journal articles. Google Scholar fits more with this kind of analysis than the Web of Science, because of a greater range of included material. Other relevant sources are field-specific databases, aggregators' and publishers' platforms[27]. The procedure of such an extensive targeted grey literature search is rather complex, far from automated and quick processing of large amounts of bibliographic information: "Organisations considered likely to be publishing materials drawing on NMDS-SC (…) were identified from existing knowledge of the sector, by the client and ourselves, and the initial findings from the bibliometric survey. A total of 24 organisational websites were manually searched and browsed, including UK government departments, sectoral bodies, knowledge intermediary organisations such as independent research organisations (…), campaigning organisations, think tanks, trade/employer organisations and the professional and mainstream press. We also conducted a limited search of social media, using social media aggregator sites" (p.297).
5. Altmetrics with grey literature produce more content but are time-consuming. Sibbald et al. (2015) conducted a case study on the inclusion of grey literature in citation analysis, based on one published article in the field of violence against women. Google Scholar and the Web of Science produced eighty journal articles citing the paper. The grey literature searches[28] found 29 other sources (27% of all results). But "this method requires additional resources. The much broader range of potential search venues demands more time and expertise. Delving into gray literature is a challenging task and requires planning and coordination, including consideration-specific inclusion/exclusion searching. Unlike database searching, common nomenclatures rarely exist for searching diverse gray literature sources; therefore, the concept of consistency in search terms across sources is difficult to achieve."

These examples confirm the potential of grey literature as source for altmetric impact assessment, with significant and complementary results based on citations, links and referrals from a broader range of scientific information, including dissertations, reports, white papers etc. But they also show that this approach is more complicated and time-consuming than the usual WoS or Scopus-based work. In contrast to traditional metrics which usually exclude grey literature, altmetrics are not limited to scientometric databases. But when it comes to larger empirical studies, this exploitation of grey literature remains an exception and is sometimes limited to science blogs, while proceedings, dissertations etc. are dismissed (see for instance Thelwall et al. 2013 or Costas et al. 2015).

Discussion

Our objective was to clarify the connection between altmetrics and grey literature. Traditional metrics have largely overlooked grey literature. Do altmetrics offer new opportunities for the development and impact of grey literature? In fact, we explored two different issues:

1. Impact assessment of grey literature – do altmetrics offer new and unique opportunities for the web-based impact measurement of reports, conferences, dissertations etc.? Do they contribute to improved visibility and impact? And if so, how?
2. Impact assessment through grey literature - how can grey literature contribute to altmetrics? Do altmetrics tools make use of grey literature?

These two issues have been described in terms of diversity and broadness, as specific benefits of altmetrics compared to traditional indicators.

[27] In this study NHS Evidence, LG Search, Hein Law Online, EBSCO Business Source Complete, Nexis and Emerald Journals

[28] Searches were conducted with Google and in Scopus, MedlinePlus, MDConsult, UpToDate, Factiva, Lexis Nexis, Google News, and Proquest Canadian Major Dailies. Major health care associations and professional organisations likely to include related content were identified, and their websites were individually searched.

Our review of recent publications, together with some altmetrics tools presents a contradictory situation:

- The potential of altmetrics for grey literature is real. Altmetric data providers like Twitter, Mendeley, Facebook or figshare but also reference managers and institutional repositories are tailored for grey literature, and they already contain significant amounts of unpublished documents.
- Assessment studies on the grey literature's web-based impact show partly higher impact than journal articles or books. Apparently, altmetrics offer a unique opportunity to exhibit the real impact of unpublished research results in conference papers, dissertations, working papers etc. and to contribute to improved visibility of these documents.
- But this work remains more or less exceptional. Most studies on F1000, Mendeley etc. include only journal articles (see for instance Mohammadi & Thelwall 2013, 2014). The main reason is that altmetrics tools need unique identifiers, standard metadata and good availability. "One of the critical issues is that these aggregators concentrate on documents that have a unique object identifier, which inevitably neglects certain document types (…) For example, Altmetric.com (…) focuses its data collection on DOIs, which has led to a de facto reduction of altmetrics studies to journal articles, excluding many types of documents and journals" (Sugimoto et al 2016).
- Impact assessment with grey literature is difficult, time-consuming and manual work, and requests expert knowledge of the scientific information landscape, especially when the grey resources are not available on open repositories but somewhere in the dark web, e.g. on less-referenced, personal or other websites.
- And then there may be other reasons to dismiss grey literature. In Hammarfelt's (2014) study research impact in the humanities, all grey items - 1,006 conference papers, dissertations and reports (20%) - were skipped from the initial corpus of 5,091 scholarly works[29] because of the "scarcity" of altmetrics data in particular for the Swedish language documents.

No identifier, lack of bibliographic control and no standard metadata, unsatisfying availability – all this is not new in the field of grey literature, and Adie's (2014) suggestions to improve the situation is only too familiar for the grey community: minimum standards for metadata (PRISM[30], DC), persistent identifiers (handle, DOI), discoverability (index, repository). His suggestion: "An open, central index of scholarly grey literature that enforced a minimum level of metadata for each item (…) An alternative would be to maintain a central index of grey literature repositories (…) and to allow harvesting from each (…)".

A central index of grey literature – this sounds like utopia. Probably the main issue is that altmetrics need DOI (Adie 2016); and the DOI appears to be the only realistic option for the assignment of permanent and citable identifiers to grey literature when it comes to prepare academic output in repositories for alternative metrics (Gerritsma 2015[31], see also Brooks & Fitz 2015). But given the history of failed initiatives for standard identifiers and metadata, we must admit that this may be just another missed opportunity.

Perspectives

Are altmetrics the future of scientometrics? For the moment, they are still "in infancy" (Erdt et al. 2016), and for many researchers, impact factor and large citation databases are still preferred for determining impact, with 'pure' altmetrics tools scoring much lower, especially in physical sciences, engineering and technology[32]. Likewise, because of not-yet achieved critical mass, lack of theory, lack of quality control mechanisms, inconsistencies and multiplicity of social web sources, data, tools and methods, Sotudeh et al. (2015) speak of

[29] Extraction from the SwePub database of academic publications at Swedish universities http://swepub.kb.se/

[30] Publisher Requirements for Industry Standard Metadata, see http://www.idealliance.org/specifications/prism-metadata-initiative

[31] Theses, working papers, reports, conference contribution – in Gerritsma's example (VU Amsterdam) grey items represents 14% of the whole output

[32] Innovations in Scholarly Communication Survey, http://altmetricsconference.com/who-is-using-altmetrics-tools/

"immaturity of the field" and call for cautious application and interpretation, even as a complement to traditional metrics. The risk of misuse and rankings based on such arbitrary information is real.

Are download counts really a metric of scholarship or only of computer activity? Is popularity an indicator of quality? How does one deal with multiple versions of the same item? For these and other reasons, Booth (2016) condemns the limited validity of the new generation of altmetrics and suspects that they follow a logic of easiness to get the data; "(they are) neither a more accurate representation of academic 'quality' nor immune to critics" (p.41). In particular, the composite, "all-in-one" Altmetric Score has been critically appraised, because of lacking of transparency, reproducibility and stability, questionable validity and significance, and problems with data sources, consistency and completeness (see Gumpenberger et al. 2016).

The "pressure of various stakeholders" and the dependency on aggregators and social media as data providers may explain one part of the criticisms (Haustein 2016). Lack of transparency and conceptual deficit are at the opposite of the purpose of the Leiden Manifesto for Research Metrics (Hicks et al. 2015) but may be related to the increasing commercial take-over of these new tools and services by those who already dominate the scientific information market.

According to Gartner's famous Hype Cycle model[33], new technology go through a typical five-phase life cycle (figure 16). After a potential technology breakthrough kicks things off ("technology trigger") and a growing number of success stories ("peak of inflated expectations") comes the "trough of disillusionment", with growing criticisms, failures and dissatisfaction.

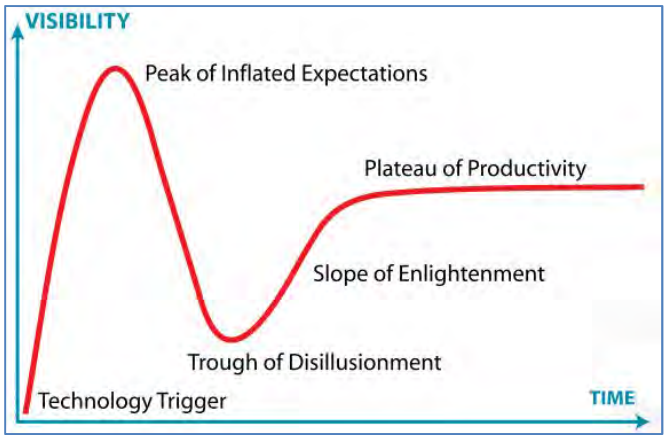


Figure 16: Gartner's Hype Cycle (source: Wikipedia[34])

On their own technology life cycle, altmetrics probably have passed by the peak of inflated expectations and are moving forward to this "trough of disillusionment", which is a necessary and salutary transition to a more realistic and satisfying situation where this new generation of metrics is no longer considered as the one and only alternative to traditional performance assessment but as new and interesting methods to assess impact of research output, complementary to traditional metrics.

Metrics shape the science, said Paul Wouters from the Centre for Science and Technology Studies at Leiden University, and we can reasonably expect that altmetrics will be part of the game. Altmetrics are already a major topic of the European Open Science Agenda and will contribute to a new rewarding and funding system.

To come back to our initial question – what is the role of grey literature in this emerging world of new assessment tools? When second- and third-generation products will appear from technology providers and later, when mainstream adoption will take off, will grey literature be part of the game or remain out of scope, just as before? For the moment, grey

[33] http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp
[34] https://en.wikipedia.org/wiki/File:Gartner_Hype_Cycle.svg

literature missed the opportunity to get on board. Since the Altmetrics Manifesto 2010, no real effort has been made to adapt the new assessment tools to grey literature or to make this literature suitable for altmetrics. Publications in the field of scientometrics show that journal (and sometimes book) publishing is still at the heart of research and development, not only for traditional metrics but also for alternative metrics. For instance, most of the contributions to the last Altmetrics Conference in Bucharest[35] are about journal publishing, and the rare exceptions deal with datasets and software, not with grey literature. Today, the future development of this new technology bears the risk of dismissing large parts of scientific literature – those parts not controlled by commercial publishers. Just as before, it is business as usual. Sometimes you don't get a second chance. But you have to be at the station when the train arrives. To get to the station means to:

- Contribute to research on altmetrics for or with grey literature, for instance in the fields of economics (working papers) or computer science (conference papers).
- Cooperate with altmetrics companies and teams for the development of appropriate tools that fit with grey literature.
- Accelerate the allocation of unique identifiers for grey literature and their authors and why not their institutions, above all this means partnership with DOI, ORCID and CASRAI[36], in particular for electronic theses and dissertations and for scientific reports.
- Contribute to further standardization of grey literature metadata.
- Contribute to increasing availability of grey literature in institutional repositories.

Getting grey literature into the heart of the coming mainstream adoption of altmetrics is essential not only for the future of grey literature in open science but also for academic and institutional control of research output and societal impact. This can be a special mission for academic librarians. Grey literature has always been a library-driven concept (Schöpfel 2010); today, as a recent survey shows, academic librarians demonstrate a higher awareness for altmetrics tools than researchers[37]. Perhaps this convergence or happy coincidence may be helpful.

[35] 3:AM Conference, Bucharest 28-29th September 2016, http://altmetricsconference.com/schedule/

[36] See the Jisc CASRAI-UK pilot on organisational identifiers https://jisccasraipilot.jiscinvolve.org/wp/2015/03/06/organisational-identifiers-working-group-outputs-and-update/

[37] Innovations in Scholarly Communication Survey, http://altmetricsconference.com/who-is-using-altmetrics-tools/ See also Malone & Burke (2016).

References

Adie, E., 2014. The grey literature from an altmetrics perspective – opportunity and challenges. *Research Trends* (37).

Adie, E., 2016. The rise of altmetrics. In: Tattersall, A. (Ed.), pp. 67-82.

Banks, M. A., de Blaaij, C., 2006. Implications of copyright evolution for the future of scholarly communication and grey literature. In: GL8 *Eighth International Conference on Grey Literature*, New Orleans, 4-5 December 2006.

Björneborn, L., Ingwersen, P., 2004. Towards a basic framework of webometrics. *Journal of the American Society for Information Science and Technology* 55 (14), 1216-1227.

Booth, A., 2016. 'Metrics of the trade': where have we come from? In: Tattersall, A. (Ed.), pp. 21-47.

Bornmann, L., 2014. Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics* 8 (4), 895-903.

Bornmann, L., Haunschild, R., Marx, W., 2016. Policy documents as sources for measuring societal impact: how often is climate change research mentioned in policy-related documents? *Scientometrics,* 1-19.

Brooks, L., Fitz, G., 2015. Grey matter(s): Embracing the publisher within. *The Foundation Review* 7 (2), 38-50.

Costas, R., Zahedi, Z., Wouters, P., 2015. Do "altmetrics" correlate with citations? Extensive comparison of Altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology* 66 (10), 2003-2019.

Dinsmore, A., Allen, L., Dolby, K., 2014. Alternative perspectives on impact: The potential of ALMs and altmetrics to inform funders about research impact. *PLOS Biol* 12 (11), e1002003+.

DORA (2012). `*San Francisco Declaration on Research Assessment*'. American Society for Cell Biology (ASCB), San Francisco.

Erdt, M., Nagarajan, A., Sin, S.-C. J., Theng, Y.-L., 2016. Altmetrics: an analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics* (first online 10 August 2016).

Galligan, F., Dyas-Correia, S., Mar. 2013. Altmetrics: Rethinking the way we measure. *Serials Review* 39 (1), 56-61.

Gerritsma, W., 2015. Altmetric opportunities for libraries. In: *2:AM Conference (Altmetrics Conference),* University Amsterdam, 8th November 2015.

Gumpenberger, C., Glänzel, W., Gorraiz, J., 2016. The ecstasy and the agony of the Altmetric score. *Scientometrics* 108 (2), 977-982.

Hammarfelt, B., 2014. Using altmetrics for assessing research impact in the humanities. *Scientometrics* 101 (2), 1419-1430.

Haustein, S., Costas, R., Larivière, V., 2015. Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. *PLoS ONE* 10 (3), e0120495+.

Haustein, S., 2016. Grand challenges in altmetrics: heterogeneity, data quality and dependencies. *Scientometrics* 108 (1), 413-423.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., Rafols, I., 2015. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520 (7548), 429-431.

Ingwersen, P., Larsen, B., Carlos Garcia-Zorita, J., Serrano-López, A., Sanz-Casado, E., 2014. Influence of proceedings papers on citation impact in seven sub-fields of sustainable energy research 2005–2011. *Scientometrics* 101 (2), 1273-1292.

Konkiel, S., 2012. Altmetrics: An app review. In: *OCLC Innovation in Libraries, post-conference event, LITA Forum 2012*, Columbus, OH, USA, October 8, 2012.

Kraker, P., Lex, E., Gorraiz, J., Gumpenberger, C., Peters, I., 2015. Research data explored II: the anatomy and reception of figshare. In: *STI2015 20th International Conference on Science and Technology Indicators*, Lugano, 4 September 2015.

Kwok, R., 2013. Research impact: Altmetrics make their mark. *Nature* 500 (7463), 491-493.

Larivière, V., Zuccala, A., Archambault, E., 2008. The declining scientific impact of theses: Implications for electronic thesis and dissertation repositories and graduate studies. *Scientometrics* 74 (1), 109-121.

Lindsay, J. M., 2016. PlumX from plum analytics: Not just altmetrics. *Journal of Electronic Resources in Medical Libraries* 13 (1), 8-17.

Malone, T., Burke, S., 2016. Academic librarians' knowledge of bibliometrics and altmetrics. *Evidence Based Library and Information Practice* 11 (3), 34+.

Moed, H. F., Halevi, G., 2014. The multidimensional assessment of scholarly research impact. *Journal of the Association for Information Science and Technology* 66 (10), 1988-2002.

Mohammadi, E., Thelwall, M., 2013. Assessing non-standard article impact using f1000 labels. *Scientometrics* 97 (2), 383-395.

Mohammadi, E., Thelwall, M., 2014. Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology* 65 (8), 1627-1638.

Neylon, C., Willmers, M., King, T., 2014. *Rethinking impact: Applying altmetrics to Southern African research.* Report, University of Cape Town. Scholarly Communication in Africa Programme, Paper 1.

NISO, 2014. *Alternative metrics initiative. Phase 1 White Paper*. National Information Standards Organization, Baltimore MD.

NISO, 2016. *Altmetrics definitions and use cases. A recommended practice of the National Information Standards Organization*. Draft. NISO RP-25-201x-1, National Information Standards Organization, Baltimore MD.

Piwowar, H., 2013. Altmetrics: Value all research products. *Nature* 493 (7431), 159.

Pontika, N., Knoth, P., Cancellieri, M., Pearce, S., 2016. Developing infrastructure to support closer collaboration of aggregators with open repositories. *LIBER Quaterly* 25 (4).

Priego, E., 2014. *On metrics and research assessment*.

Priem, J., Taraborelli, D., Groth, P., Neylon, C., 2010. *Altmetrics: A Manifesto*.

Priem, J., Hemminger, B. H., 2010. Scientometrics 2.0: New metrics of scholarly impact on the social web. *First Monday* 15 (7).

Prost, H., Le Bescond, I., Schöpfel, J., 2010. Usage assessment of an institutional repository: a case study. In: *GL12 Twelfth International Conference on Grey Literature*, Prague, 6-7 December 2010.

Schöpfel, J., 2010. Towards a Prague definition of grey literature. In: *GL12 Twelfth International Conference on Grey Literature*. Prague, 6-7 December 2010

Schöpfel, J., Prost, H., 2009. Usage of grey literature in open archives: state of the art and empirical results. In: *GL11 Eleventh International Conference on Grey Literature*. Washington D.C., 14-15 December 2009.

Sibbald, S. L., MacGregor, J. C. D., Surmacz, M., Wathen, C. N., 2015. Into the gray: a modified approach to citation analysis to better understand research impact. *Journal of the Medical Library Association: JMLA* 103 (1), 49-54.

Sotudeh, H., Mazarei, Z., Mirzabeigi, M., 2015. CiteULike bookmarks are correlated to citations at journal and author levels in library and information science. *Scientometrics* 105 (3), 2237-2248.

Sud, P., Thelwall, M., 2014. Evaluating altmetrics. *Scientometrics* 98 (2), 1131-1143.

Sugimoto, C. R., Work, S., Larivière, V., Haustein, S., submitted. Scholarly use of social media and altmetrics: a review of the literature. *Journal of the American Society for Information Science and Technology*.

Tattersall, A. (Ed.), 2016. *Altmetrics: a practical guide for librarians, researchers and academics*. Facet Publishing, London.

Thelwall, M., Haustein, S., Larivière, V., Sugimoto, C. R., 2013. Do altmetrics work? Twitter and ten other social web services. *PLOS ONE* 8 (5), e64841+.

Urquhart, C., Dunn, S., 2013. A bibliometric approach demonstrates the impact of a social care data set on research and policy. *Health Information & Libraries Journal* 30 (4), 294-302.

Wilkinson, D., Sud, P., Thelwall, M., 2014. Substance without citation: evaluating the online impact of grey literature. *Scientometrics* 98 (2), 797-806.

Wouters, P., Costas, R., 2012. Users, narcissism and Control—Tracking the impact of scholarly publications in the 21st century. In: *STI2012 17th International Conference on Science and Technology Indicators (STI)*, 5-8 September, 2012, Montreal, Quebec, Canada.

Yu, M.-C., Wu, Y.-C. J., Alhalabi, W., Kao, H.-Y., Wu, W.-H., 2016. ResearchGate: An effective altmetric indicator for active researchers? *Computers in Human Behavior* 55, 1001-1006.

Zoller, D., Doerfel, S., Jäschke, R., Stumme, G., Hotho, A., 2016. Posted, visited, exported: Altmetrics in the social tagging system BibSonomy. *Journal of Informetrics* 10 (3), 732-749.

The complete bibliography with links and further readings is available at http://www.citeulike.org/user/Schopfel/tag/gl18

All websites visited in August and September 2016.

Acknowledgment to Paul Needham, JISC, for helpful advice with the IRUS-UK project.

SUBSCRIBE, SEARCH, DISCOVER AT:
Greylit.org or greylithelp@nyam.org

Enhance the impact of your research with contemporary, comprehensive urban health resources, selected from more than 750 non-commercial publishers from around the world.

- Researched and curated to give you a unique, inside track on urban health issues.
- Indexed and described to make it easy to discover resources across multiple disciplines.

Uncovering hidden resources for researchers, health professionals, librarians and the public.

NYAM.org **The New York Academy of Medicine** 1216 Fifth Avenue | New York, NY 10029

The GreyLit Report: Understanding the Challenges of Finding Grey Literature

Danielle Aloia and Robin Naughton,
The New York Academy of Medicine Library, USA

Abstract
Searching for and finding grey literature can be difficult. Grey literature, by its nature, is not commercially published and as a result, it requires multiple search strategies to identify and curate quality literature on a subject. Our study into how researchers share grey literature (Aloia and Naughton, 2015) found that researchers speak with colleagues, subscribe to listservs/newsletters, and go to organization websites to find current grey literature. In order to better understand the needs of the health sciences research community, we interviewed GreyLit Report users about their challenges, tools and methods for finding grey literature. The Grey Literature Report (GreyLit Report), developed in 1999 by The New York Academy of Medicine, is a centralized location that makes it easier for health researchers to find grey literature in their field. Speaking directly to librarians and researchers about their needs helped us to better understand how the GreyLit Report website can be enhanced to respond to those needs. Over the course of a week, we conducted online interviews with national and international users of the GreyLit Report. Based on this study, the researchers learned how the GreyLit Report can be enhanced to better serve the grey literature community and add to the growing need for a centralized location to find grey literature. In addition, the paper provides a template for planning and reporting of grey literature searches based on extensive analysis of the research literature.

Introduction
Commercial databases are abundant, structured and time-honored tools that are the go-to source for researchers and librarians looking for quality resources. This makes it easy for researchers to document and capture the process used to find relevant articles, potentially making the search process reproducible to others. Searching for peer-reviewed articles is standard for the systematic review search process and commercial databases are structured and indexed in a way that facilitates effective searching. But as more research is published through alternative channels, the variability in search strategy has grown. Oftentimes, these materials are published on organization websites or within repositories that are structured and indexed in multiple ways. Organization websites are extremely variable and cannot be counted on to find relevant results. On the other hand, repositories maintained by a university or specific academic discipline can be tailored to meet the needs of their respective users. For these reasons, there is no one place or systematic process to search for grey literature, and as such grey literature searching requires a different set of skills.
This research study seeks to understand how researchers and librarians search for grey literature and what resources they use. To do this, the study included a detailed literature review analyzing health science literature that used grey literature for systematic review searching, and semi-structured interviews with researchers and librarians who search for grey literature. The study was guided by three research questions.

Research Questions
1. What challenges do researchers and librarians face when trying to find grey literature on the health sciences?
2. How are researchers and librarians in the health sciences searching for grey literature?
3. To what extent does the Grey Literature Report help researchers find grey literature in their field?

Literature Review
Much has been written about the use of grey literature in the research process and its importance in the systematic review but less attention has been given to guidelines concerning the search process and reporting for grey literature. For instance, the *Cochrane*

Handbook for Systematic Review of Interventions 2011, provides a small paragraph on grey literature and mentions three or four sources, but no search techniques. The Canadian Agency for Drugs and Technologies in Health (CADTH) publishes *Grey Matters* and *Methods and Guidelines series* to alleviate the myths behind searching for grey literature. Their publications provide excellent guidance on search techniques in health technology assessment that can be used as a model for other topical searches. In *Methods Guide for Medical Test Reviews,* the U.S Agency for Healthcare Research and Quality suggests "Combining highly sensitive searches utilizing textwords with hand searching and acquisition and review of cited references in relevant papers is currently the best way to identify all or most relevant studies for a systematic review." Currently, there is no one set of search techniques available for conducting a grey literature search. Grey literature searching may be systematic but it is time consuming, hard to replicate or reproduce, and resources vary among disciplines (Adams, 2016). As a result, this literature review explored the tools researchers used to search for grey literature, how search techniques and results were reported, and what types of search strategies were executed.

Methods

The literature review is based on research papers collected from a PubMed search. Articles and reports that focused on systematic review research were collected and analyzed. Criteria for inclusion included English language, review article/report, mention of search strategies, and published in the past 5 years. The PubMed search (("grey literature" OR "gray literature") Limits: English, review, and within 5 years) yielded over 1500 results. These results were paired down to 400. Of these 103 were excluded for various reasons, as listed below:

- Fifty-seven focused on grey literature search strategies for systematic reviews,
- Twenty-one didn't report how or where they searched, so we were unable to determine the search methodology,
- Twelve didn't report how or where they searched,
- Twelve papers indicated they just searched the traditional databases PubMed, Web of Science to locate grey literature,
- Six we were unable to obtain the full text of the article,
- Five of the articles only reported on the selection bias of not including grey literature in the systematic review, and
- Three papers mentioned explicitly that they excluded including grey literature in their searches.

A total of 297 papers were analyzed for the types of sources reported for searching grey literature. The 57 articles specifically discussing search strategies related to grey literature searching were analyzed for consistency of search methodology and to develop a model to conducting a systematic search.

Reported Tools for Grey Literature Searching

Systematic review methodologies require the use of grey literature to overcome bias and to be sure that all relevant literature is captured. In the 297 papers analyzed, the most common search tools reported for searching grey literature were websites, handsearch, and Google. Clinical trial databases and registries were the most highly cited (155 mentions) but were not included in this analysis because it is a specific type of grey literature and has its own resources. In Figure 1, the number of mentions means that most authors reported using more than one tool in their search strategy.

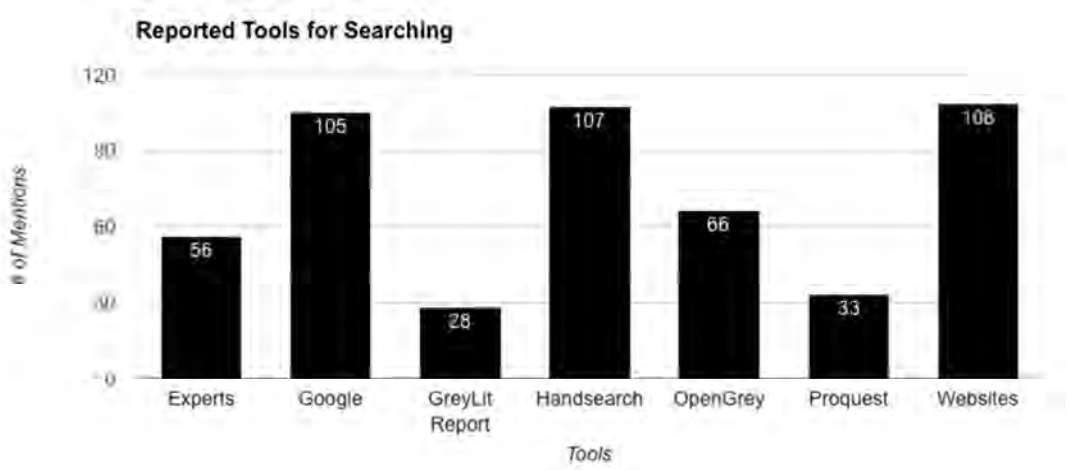


Figure 1: Reported Tools for Searching

OpenGrey, ProQuest, and the GreyLit Report are specialty databases and were cited by about 20% of the papers. Contacting experts to find relevant research was cited highly (56 mentions) as well and is a time honored method within the systematic review process. Handsearching is the "gold standard" for validating peer-review search results and making sure all relevant articles on the research topic are captured. This is almost impossible to do for grey literature searching therefore making it difficult to validate (Adams, 2016). Since Google and websites are the most cited sources for searching for grey literature it is essential to have some standard guidelines for searching. This will help make the search results and reporting transparent and maybe even reproducible.

Search Strategies

Of the 400 reviews that were analyzed, 57 were specific on how to search for grey literature for systematic reviews. Each article recommended a different set of search strategies for a variety of platforms. A key insight was that a good grey literature search includes these resources: online databases, websites and search engines, repositories, online catalog, and asking experts (Mahood, 2014). This is not surprising because these sources have been highly cited in the literature. Grey literature searching can be time consuming and each grey literature source can take up to 1.5 hrs to search (Saleh, 2014). A good Google search strategy is to use advanced search techniques by searching the title field and checking results past the first 10 pages (Haddaway, 2015). Stansfield, et al. (2016) recommend putting effort into planning the search process early, especially, when it comes to locating resources to search.

Search Reporting

Of the 400 papers collected 21 did not report any search strategies in their results. A few listed all the resources searched without keywords, some listed resources with keywords used, or included the results as an appendix or supplement. Briscoe (2015), *Web Searching for Systematic Reviews,* analyzed search results of systematic review papers and found that, the majority of papers only reported one search detail and few papers provided links to resources. Briscoe (2015) provides recommendations on reporting results for websites and search engines. Godin, et al. (2015), *Applying Systematic Search Strategies to Grey Literature*, provides guidelines on how to perform and record a grey literature search. Godin et al. provide an amazing amount of detail about their search methods and results in a short and easy to read paper! They created a multi-search plan using four resources: grey literature databases, Google, websites, and experts. They dedicate a paragraph to each search strategy and include an outline of the results in a PRISMA diagram. The authors also provide a timeline for the different methods of their search plan and attach additional files of their search results for each resource.

Below is a checklist of items to include in the search process. For best results in recording and tracking your searches apply these elements to your search strategy:

- Provide name of organization/search engine
- URL
- Dates searched

- Search terms or keywords used
- Some analysis of the results
- How and why results were chosen with links to selected sources
- Time spent

In-Person Interviews
Interviews were conducted to understand the needs of GreyLit Report users in relation to how they search and their experiences using the GreyLit Report website.

Methods
Using qualitative research methods with a focus on the semi-structured interview, participants were recruited and interviewed about their experience with the GreyLit Report.

Recruitment
Participants for the research study were recruited from the GreyLit Report subscribers. There were more than 2000 subscribers to the GreyLit Report who were contacted through email regarding the study. An original email and follow-up emails were sent to subscribers describing the project. Participants were also recruited from an earlier study where they were asked to provide contact information if they would like to be interviewed regarding the GreyLit Report.

Participants
There were a total of 12 interviews with 14 participants. Two interviews were group interviews because participants invited colleagues to join the scheduled interview who were then included in the discussion. Eleven participants were females and three participants were male. The participants represented a global reach from the United States of America, Australia, Canada, Netherlands, and the United Kingdom. There were librarians, information professionals and researchers. Table 1 shows the breakdown of participants by location and job title.

Table 1: Participants based on location and job title

Location	**Job Title**	**# of Participants**
Australia	Liaison Librarian	1
	Librarian	2
	Teaching and Learning Librarian	1
Canada	Information Specialist	1
	Manager, Information Services	1
Netherlands	Director	1
United Kingdom	Information Officer	1
United States of America	Head of Information Services	1
	Graduate Student	1
	Graduate Student/professor	1
	User Services Librarian	1
United States	Informationist	1
	Research Librarian	1
	Senior Research Advisor (Australia)	1
Total		**14**

Data Collection & Procedure

Data was collected using semi-structured interview protocol. The interview guide was divided into three parts. The first part of the interview asked participants about their overall experience with grey literature and how they search for grey literature. The second part of the interview focused on the GreyLit Report and asked participants about their experience with the GreyLit website. The third and final part of the interview asked participants demographic information regarding job titles, gender and age range.

All interviews were remote interviews, conducted using either phone or Skype (an online remote conference tool) and lasted for one hour. Participants provided their preferred method of contact and the researchers accommodated them accordingly. Using an audio recorder, interviews were recorded for later transcription and analysis. The interviews were transcribed and analyzed for major themes related to the research questions and study goals.

Data Analysis

Standard thematic analysis (Wildemuth, 2009) was used to analyze the qualitative data. The process includes repeated readings to define and refine the themes and categories throughout the data. This repetition is necessary to identify the categories and refine the coding scheme for the data in order to answer the research questions.

Results

Search & Databases Used

Participants were asked how they searched for grey literature and what tools/databases they used for search. Each participant provided multiple responses. Results varied in terms of search strategies and goals, but the majority of participants used Google (10) and organization and government websites (11) to search for grey literature, and OpenGrey (6).

Figure 2: Word Cloud of Participants Search

Participants suggested that selecting search tools depended on the strategies and goals of the search in the research topic. Systematic review research was discussed as one of the reasons participants searched for grey literature.

When asked about specific databases, participants mentioned over 69 databases used, including commercial databases such as PubMed, EBSCO, Web of Science, Proquest Scopus, country-specific databases such as Australian Policy Online and DANS Archive, topic-specific such as Clinical Practice Guidelines, and many more databases particular to the participants' area.

Topical Searches of Interest

Participants were asked what topical searches would be of interest to them and their domain. They reported over 75 concrete responses in the areas of health information technology, public health, clinical medicine, legal, and physical exercise.

Other Resources Indexed
Participants were asked what other types of grey literature they would like to see indexed in the GreyLit Report. Eight participants mentioned Conference Proceedings, three mentioned Datasets, and three mentioned Webinars.

Experience with the GreyLit Report
Participants were asked about their experience with the GreyLit Report. In half of the interviews (6), participants used the GreyLit Report as a current awareness tool and would recommend it to others searching for grey literature. In some interviews (5), participants stated that they had a good experience with the GreyLit Report website. Based on the data, a majority of participants did not have as good an experience with the website and suggested that a few aspects of the site could be improved, including clarity on the Academy priority areas, an improved search, and broader scope or coverage.

Discussion

What challenges do researchers and librarians face when trying to find grey literature on the health sciences?
The major challenge faced when trying to find grey literature is the variety of sources available among and between disciplines. Respondents to the interviews indicated that sources they used depended upon the research question. Another challenge they faced was that each source has its own search criteria, making it hard to be consistent with search terminology.

How are researchers and librarians in the health sciences searching for grey literature?
Librarians and researchers are using a variety of ways to search for grey literature. The most common method, besides the traditional databases, are Google and organizations websites. This was found to be true both in the literature and from in-person interviews.

To what extent does the Grey Literature Report help researchers find grey literature in their field?
The Grey Literature Report was helpful to users in that it provided access to publisher names and alerts to new and current resources. Respondents reported difficulty with the search functionality and didn't feel the scope of content was broad enough.

Conclusion
Researchers and librarians use a variety of methods to find grey literature. This is supported by both the research literature and the interviews in this study. Google and organization websites were the most cited sources to find grey literature, followed by contacting experts and handsearching the peer-reviewed literature. The GreyLit Report was the fifth most cited in the literature and highly recommended by librarians as a good source for grey literature. The study showed that the major challenge of finding grey literature has to do with the nature of grey literature itself. The same methods used to search commercially published resources cannot be used to find grey literature resources. Grey literature requires a different search strategy and plan.
We recommend developing a search plan for finding grey literature, decide on the resources to use before starting your search, indicate search strategies for each resource using the guidelines above, and make note of results. Be prepared to spend time on each step of the process.

Bibliography

Adams, Jean, Frances C. Hillier-Brown, Helen J. Moore, Amelia A. Lake, Vera Araujo-Soares, Martin White, and Carolyn Summerbell. "Searching and Synthesising 'Grey Literature' and 'Grey Information' in Public Health: Critical Reflections on Three Case Studies." *Systematic Reviews* 5, no. 1 (2016): 164. doi:10.1186/s13643-016-0337-y.

Aloia, Danielle, and Robin Naughton. "Share #GreyLit: Using Social Media to Communicate Grey Literature." *The Grey Journal (TGJ): An international Journal on Grey Literature* v. 12, no. 2 (2016): http://hdl.handle.net/10068/1024651.

Briscoe, Simon. "Web Searching for Systematic Reviews: A Case Study of Reporting Standards in the UK Health Technology Assessment Programme." *BMC Research Notes* 8 (2015): 153. doi:10.1186/s13104-015-1079-y.

CADTH Information Services. "Grey Matters: A Practical Tool for Searching Health-Related Grey Literature." Canadian Agency for Drugs and Technologies in Health, 2015. https://www.cadth.ca/resources/finding-evidence/grey-matters.

Godin, Katelyn, Jackie Stapleton, Sharon I. Kirkpatrick, Rhona M. Hanning, and Scott T. Leatherdale. "Applying Systematic Review Search Methods to the Grey Literature: A Case Study Examining Guidelines for School-Based Breakfast Programs in Canada." *Systematic Reviews* 4 (2015): 138. doi:10.1186/s13643-015-0125-0.

Haddaway, Neal Robert, Alexandra Mary Collins, Deborah Coughlin, and Stuart Kirk. "The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching." *PloS One* 10, no. 9 (2015): e0138237. doi:10.1371/journal.pone.0138237.

Mahood, Quenby, Dwayne Van Eerd, and Emma Irvin. "Searching for Grey Literature for Systematic Reviews: Challenges and Benefits." *Research Synthesis Methods* 5, no. 3 (September 2014): 221–234. doi:10.1002/jrsm.1106.

Saleh, Ahlam A., Melissa A. Ratajeski, and Marnie Bertolet. "Grey Literature Searching for Health Sciences Systematic Reviews: A Prospective Study of Time Spent and Resources Utilized." *Evidence Based Library and Information Practice* 9, no. 3 (2014): 28–50.

Stansfield, Claire, Kelly Dickson, and Mukdarut Bangpan. "Exploring Issues in the Conduct of Website Searching and Other Online Sources for Systematic Reviews: How Can We Be Systematic?" *Systematic Reviews* 5 (November 2016). http://download.springer.com/static/pdf/34/art%253A10.1186%252Fs13643-016-0371-9.pdf?originUrl=http%3A%2F%2Fsystematicreviewsjournal.biomedcentral.com%2Farticle%2F10.1186%2Fs13643-016-0371-9&token2=exp=1479485891 acl=%2Fstatic%2Fpdf%2F34%2Fart%25253A10.1186%25252Fs13643-016-0371-9.pdf* hmac=dc05b476c34a03e4497ac097310af0f9154dbc6efd5135366638c5790f59a6f9.

Wildemuth, Barbara M. *Applications of Social Research Methods to Questions in Information and Library Science.* (2009). Westport, CT: Libraries Unlimited.

Korea Institute of Science and Technology Information (KISTI)

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

*** Vision**

World-class information research institute creating values for customers

*** Main functions**

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

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

***Contact information**

KISTI email address: hcpark@kisti.re.kr

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

Debate about Scientific Popularization in Russian Public Sphere (Based on Grey Literature Material)

Yuliya B. Balashova
Saint Petersburg State University, Russia

Abstract
The article is devoted to the problematic field associated with the popularization of science, in the reflection of the grey literature. In Russia, the public sphere is arranged in such a way that still many important issues do not discussed in the press, but on various discussions, which materials fixed in the grey literature.
Keywords: grey literature, popular science journalism, debate about scientific popularization, Russian public sphere.

Introduction
At the present time, in Russian increased public interest to scientific problematic. Knowledge itself is a value in all times. Traditionally, science held a high position in Russian society, was vastly included into the public sphere. In Soviet times, the unique model of communication between science and society has been designed, and it didn't duplicate Western experience. Now in Russia a significant number of different activities, related to the science promotion, begin to be carried out. Different conferences, public discussions host in different regions of the country. Media begin to show more interest in all issues of science communication. However, in most cases, evidences of this discussion most accumulated not in the media, but in the grey literature. This is a very interesting phenomenon, peculiar exactly to the Russian public sphere. Traditionally, the most important questions of the public life were not fully reflected in the media. Historically, they informally discussed in friendly circles, clubs, recorded in protocols of the various societies, and in private correspondence (Aronson, M., and Reyser, S., 2000). These documents can be considered as prototype of the "grey literature". Even in the middle of the 19th century in Russia verbal social communication is still in many cases prevailed over writing (Pirozhkova, T. F., 1997). This trend continued in the conditions of Soviet life. In the late Soviet period, the most pressing social issues were discussed mostly in the locker rooms, at the kitchens in the apartments. But in Soviet times, science had a great reputation in the society and seriously supported by the state.

In the modern Russia, science became one of the most important national priorities again. The process of science popularization was broken in the post-Soviet time. As a result, in this area have been accumulated a lot of problems. One of the most significant problems lies in the fact that science in Russia is not sufficiently mediatizated. Therefore, actual discussion about the ways of science development is mostly accumulated in grey literature.

Organisation and methodology of research
The main method of research is the method of participant observation. The author engaged into the process of science popularization in the aspects of education, research, as well as in the public activities. I am myself the creator and the head of the new master's program "Popular science journalism", which started two years ago at Saint Petersburg State University, Russia. The author also is the organizer and participant in a number of scientific conferences, discussions, seminars, devoted to science communications, and science promoting. There materials are mainly reflected in the conference programs and post-releases. Accordingly, they are available to a relatively small circle of interested parties. This is the main reason why this discussion is centered around the same issues forming a germenevtic circle.

The debates basically come down to two main issues that have a pronounced methodological nature.

The first of them is the following question: are journalists able to popularize science? A positive answer to this question has long been known. However, here is occurred

polarization of the points of view. Russian scientists believe that science journalists and writers are only make harm (Ivanov, I., 2007). Representatives of the communicative sphere are confident that Russian scientists themselves at the moment are not able to explain to the public the essence of their work (Vaganov, A. G., 2007).

Model of the knowledge in the West is based on the separation of science / art. In such logocentric country, like Russia, this dichotomy is not entirely justified. The very type of national consciousness tends to traditionalism and syncretism. In addition, Russia has not had the historical preconditions for the formation of narrow specialization. One of the most significant cultural reasons is not as consistent as compared with the Western Europe a Russian classic hierarchy development. The idea of the commonwealth of sciences is central to the whole tradition of the national popular science journalism. In the classic Russian "thick" magazines: "Contemporary" ("Sovremennik"), "Notes of the Fatherland" ("Otechestvennyie zapiski") departments of science and literature were mixed. The first issue of the Russian brand of popular science journalism was published in 1890. We are talking about the magazine "Science and life" ("Nauka i zhiz'n"), positioned itself as a "literary, artistic, social, and popular-science magazine". All subject areas were representative in terms of cognition; pre-revolutionary "Nauka i zhiz'n" was opened by mixed department "Science and art". In the popular science magazine science was presented as knowledge of the whole world. Approach to the understanding of scientific knowledge as universal knowledge was very typical to the classical Russian popular science magazine.

Another issue causing debate: is it necessary to popularize humanities? In modern Russian scientific environment has developed evaluative attitude towards humanitarian knowledge as opposed to natural science. In this sense, a logical continuation should be implicit question: "Is humanities research?". Russian grey literature paradoxically convinced that only natural and technical knowledge is a true science. So, here is a characteristic split in modern Russia between humanitarian and natural science. However, exactly humanitarian component has been put into the base of the Russian system of knowledge publicity. Historically has been accumulated considerable experience in the public implementation of humanities: from scientific and popular publications to educational films and lectures (Balashova, Yu. B., 2014).

The main body of the Russian grey literature related to the topic of pseudoscience and the fight against pseudoscience. So, in 1998 on the base of the Russian Academy of Sciences was created the special "Commission Against Pseudoscience and Falsification of Scientific Research". It has a coordination character, also publishes an annual bulletins "In Defense of Science", and conducts a variety of scientific and educational activities. But we have to note that passion for combating pseudoscience appears to be the legacy of the Soviet exposing companies. The latest trend of struggle against pseudoscience, in our view, does not introduce new meanings in a conversation about science popularization, as calls for a system of prohibitive measures, and based on the rhetoric of annihilation.

Discussion

In those cases, when discussion about science popularization proceeds to the media, it does not become more meaningful because it not bases on the previous experience, reflected in the grey literature. As an example, we give a representative event, which was held on June 28, 2016 in the upper house of the Russian parliament – the Federation Council.

It had potential of a large conference and represented a constructive attempt to combine different paticipants of the objectively difficult process of scientific popularization. Among them were representatives of the university and academic community, science journalists. Within the framework of this debate it was able to overcome the estimated range of conflicting judgments accompanying discussion around the problem field: science – journalism.

During the speeches sounded the idea that society needs a popular science again. The authority of the Russian science remains the highest in the world, but it lacks the publicity. Therefore, the Russian scientific sphere in particular, needs a mediator between knowledge

and society. In the developed in science promotion countries, and especially the United States, the ability to external communications is competence of the modern scientist, and any academic institution is accompanied any serious research by information campaign.

Despite the constructive nature of the discussions, this event caused a negative response in the press. The author of the negative response in the leading popular science Russian newspaper “Troitsky variant” (under the name of Moscow suburb) was a student (or may be – gratuate stident) (Troitsky variant, 2016). In accordance with her position, the government is not able to organize this kind of events, as, indeed, to engage science popularization. This position, in principle, contrary to all history and practice of science popularization in Russia. In addition, this view completely ignores the experience of similar events, reflected in grey literature.

From my side, I acted in the same newspaper with an alternative evaluation (Troitsky variant, 2016). And this position received support in the other media (Ecology and life, 2016).

If the discussion on the issues about science popularization takes into account a variety of materials published in grey literature, it would be much more constructive.

Conclusion

Materials of the different conferences and events, discussion in the blogosphere make up the range of sources about science popularization that require reflection. The scientific studies of the nature, content, specific aspects of the science popularization in modern Russia practically not involved.

Meanwhile, the controversy in the public space towards the popularization of science is centred on those issues that have long known. In the field of public actively discussed the issue could journalist popularize science or not. In Soviet times, the objective of which was to raise the level of the mass audience to scientific, science successfully popularized as the scientists themselves, as well as journalists. A similar situation typical for the Western press too.

Russian public sphere as a whole is arranged in such a way that in order to create a more holistic view of what is happening, it's necessary to appeal to grey literature.

Acknowledgment

The authors gratefully appreciate the support of the Russian Humanitarian Scientific Fund.

Funding Information

The article has been prepared with the financial support of the Russian Humanitarian Scientific Fund (RHSF). The project No. 16-03-50128.

Ethics

This article is original and contains unpublished material. The authors confirm that there are no ethical issues involved.

References

Aronson, M., and Reyser, S. (2000), Literary circles and salons. St. Petersburg: Academic project, 400 p.

Balashova, Yu. B. (2014), General scientific foundation of the new master's program “Popular scientific journalism”. – In: Pseudoscience in the modern world: Mediasphere, higher education, school. St. Petersburg: Publisher VVM, pp. 21 – 25.

Ecology and life, 2016. – http://ecolife.ru/zhurnal/articles/43490/

Ivanov, I. (2007), Anatomy of updates, or how physics actually study the elementary particles. – http://elementy.ru/nauchno-populyarnaya_biblioteka/430431/Anatomiya_odnoy_novosti_ili_Kak_na_samom_dele_fiziki_izuchayut_elementarnye_chastitsy.

Pirozhkova, T. F. (1997), Slavophil journalism. Moscow: Moscow State University Press, 221 p.

Troitsky variant, 2016. – http://trv-science.ru/2016/07/26/nauchnaya-zhurnalistika-na-ploshhadke-soveta-federacii/#comments

Vaganov, A. G. (2007), Popular science journalism and prestige of science in the public consciousness. – In: Russian Chemical Journal. 51, pp. 86 – 90..

'Grey crossroads' in cultural heritage preservation and resource management

Luisa De Biagi, CNR Central Library
Roberto Puccinelli, Telecommunications and Informative Systems Office, Italy

Abstract
Among the assets that make up the cultural heritage of a country, a special place is assigned to the internal documents produced by organs and entities belonging to the Public Administration. In the public sector, for example, the minutes of meetings of the Boards of Directors are considered historical documents and as such are preserved in the for a long time. Actually from them it is possible to gain insight about the genesis of important decisions which affected the lives of many people. In some countries there is a legal obligation to deposit those documents in long term digital preservation systems, which adhere to ad hoc defined standards. In our opinion, many of those documents can be reckoned as grey literature assets and, beyond "plain and simple" preservation, some additional measures may be deployed in order to extract information and insights from them. In this paper we illustrate a process to collect those assets, cleanse and enrich their metadata and then store them in ad hoc defined data marts, upon which Business Intelligence tools can be used for data navigation and analysis. We finally show some examples of insights that may be acquired from such analysis.

Introduction

Grey literature is a field in library and information science that deals with the production, distribution and access to multiple document types produced on all levels of government, academics, business, and organization in electronic and print formats not controlled by commercial publishing, i.e. where publishing is not the primary activity of the producing body (Greynet, 2011).

Grey literature is produced by entities whose main goal is not publishing. The grey material produced by public administrations and public/ industrial research laboratories, is significant both quantitatively and qualitatively, but has restricted dissemination for internal use. It should be noted that organizations, researchers and academics supplying scientific and technical information in grey literature, trust and support this type of documentation, not only because it contains more detailed outcomes and data, but also because it's produced up to 12 or even 18 months before the official papers are published.

Thus, the two way problem is: accessing grey literature for researchers and reaching identification and acquisition of grey literature for librarians and other information professionals. Moreover, the impact of Grey literature is largely dependent on research field/disciplines and its subjects categories, on methodologies approaches and on sources used. Nowadays newsletters, e-mails, blogs and other social networking sites are community based kinds of GL.

Practice Guidelines are highly important to biomedicine and nursing, working papers are used in Social Sciences (particularly Economics) and patents are important for the so called "hard sciences" (Physics, Geophysics, Chemistry, Biology, Maths) and for applied/technical sciences as Engineering.

Research data are also a kind of grey literature, even considering Social Sciences and Humanities: for example census, geospatial and economic data are used by local governments to formulate policies about preservation, valorization and risk evaluation.

Actually, other branches of cultural heritage are focused on technical aspects, such as archeology and archeometrics, architecture, diagnostics, preservation and restoration, also used for risk evaluation (e.g.: see recent Italian earthquakes and the current discussion about restoration and rebuilding policies and plans).

By consequence, given that some results remain hidden, not published anywhere but internal 'grey'sheets or data-set archive, there might be more integration between the digital preservation and the various commercial tools and settings managing the long term accessibility of records in database systems useful to decision makers (data warehousing).

Another interesting source of grey literature is the administrative activity of organizations. Many documents are produced during the lives of organizations that in some way trace their history, evolution and end, sometimes providing detailed information about the motivations behind events and decisions. In many countries, there is currently a legal obligation regarding long-term digital preservation of this type of documents. This means not only that the assets will be preserved, but also that their long-term readability will be guaranteed. Moreover, this provides a chance for collecting and controlling useful metadata about the above mentioned assets.

One interesting example are the minutes of the Boards of Directors' meetings. They are considered historical documents and as such must be preserved in the long run. Actually, from them it is possible to gain insight about the genesis of important decisions, which affected the lives of many people.

We think that this long-term accumulation of documents and data lays the basis for the implementation of data warehouses, upon which many types of analysis could be performed. To achieve that goal, a well-designed process must be put in place in order to collect, cleanse and enrich high quality contents and metadata.

In the following we present a short description of models, workflows and architectures for digital preservation. We then describe Business Intelligence framework for metadata and content analysis. We finally illustrate some best practices and European project examples.

Digital Preservation: models, workflows and architectures

Digital Preservation can be considered a mature field in which a well identified set of standards and best practices has been developed over time and accepted by the reference community.

The **Open Archival Information System (OAIS) model** is currently considered the reference model for Digital Preservation Systems. It has been developed in the context of space agencies, which needed to preserve the huge mass of data coming from satellite observations. First introduced as recommendation in 2002, it has been promoted to ISO standard in 2003 (**ISO 14721**), subsequently updated in 2012 (**ISO 14721:2012**).

It does not describe a technical architecture nor it assumes a particular one. It just defines:

- the **functions** a DP system should implement,
- the **actors** that interact with the system (content producers, repository managers, content consumers),
- the supported **workflows** and
- the **digital objects** that are stored, managed and exchanged with the external world.

According to the OAIS model, a DP system should implement a set of functions that can be associated to the following functional entities:

Administration: allows the configuration of the system and the coordination of all the other entities;

Ingest: accepts contents and related metadata;

Archival storage: stores contents and related metadata and provides access to them;

Data management: collects and maintains the information required for the management and access of the contents;

Preservation Planning: monitors the contents and guaranties the access and readability of the contents for the designated community, also executing format migration in case of obsolescence;

Access: provides controlled access to contents through Authentication and Authorization mechanisms.

The elementary unit for a DP system is the **Information Object (IO)**, which comprises the **Data Object** (**DO** - the actual digital content, represented as a bit-stream conforming to a particular format) and the **Representation Information** (**RI**) needed for the correct fruition and interpretation of the digital content. A DP system accepts, manages and generates **Information Packages (IP)**, which are composed of IOs. The OAIS model identifies the following three types of IP, each one containing one or more digital contents and the related meta-information:

- **Submission Information Package (SIP)**, used by the content providers to submit their objects to the DP system,
- **Archival Information Package (AIP)**, used by the system to archive contents (the contents provided in one SIP could be store in one or model AIPs),
- **Dissemination Information Packages (DIP)**, composed by the system in response to the access requests submitted by consumers.

Let's take the case of AIPs, which are the ones actually stored in a DP system. They are composed of **Content Information** (generally implemented as an IO containing the actual digital content and the related metadata) and the **Preservation Description Information (PDI)**, which is composed of one or more IOs that provide the following types of information:

Reference Information, which regards the identifiers assigned to the digital content;

Context Information, which documents the creation context of the digital content;

Provenance Information, which describe the origin and the history of the object;

Fixity Information, which is used to certify provenance and integrity of the object.

In a Digital Preservation system, knowledge can be extracted both from contents and metadata, the latter being the easiest to leverage, being by nature structured and (hopefully) controlled. The main standards used in DP systems for packages and metadata formats are:

- **UNI-SInCRO 11386: 2010**, for IP's structure (with particular reference to AIPs),
- **ISO 15836:2003** Information and documentation - The **Dublin Core** metadata element set, for metadata in general.

Analysis workflows

In the present section we describe a viable workflow that allows the creation of grey literature Data Warehouse. The reference scenario features a set of repositories (belonging to one or more organizations) that represent the main source of information for the whole system. The workflow can be divided into two main streams: one for metadata the other for actual contents.

Metadata

Metadata can be extracted from the system using ETL procedures (Extract-Transform-Load) that fetch metadata from the repositories, perform cleansing, enrich them using of external data sources and reconcile them in a unified data base structure. Additional ETL procedures are then used to populate ad hoc data marts designed to address different types of inquiries and users. Data visualizations and analysis tools are then used to create queries and display results by means of tables and charts.

Textual contents

Textual contents are still a challenging asset to manage where it comes to knowledge extraction, due to their unstructured and often inconsistent nature. There are two interesting approaches to extract information from textual unstructured documents. Nevertheless, there is a wealth of text processing techniques and algorithms that can help getting insights from document bases. We identify two interesting approaches in this field: one aimed processing unstructured data to extract knowledge that can be expressed in terms of concepts, relations, topics, categories, etc., which are themselves represented by text; the second is focused on extracting textual and numerical information from

documents, that can be fed into a classic data warehouse and used to produce tables and charts (i.e. provide quantitative insight). The latter approach is closer to classic decision support systems, but the former can provide useful, non-trivial insight. To facilitate the following exposition, we shall call the first approach "semantic" and the second "quantitative". These terms do not match with any standard or universally accepted taxonomy and will be used just for the sake of clarity.
Examples of the first type of analysis are:

- **Information extraction**
  These techniques allow the extraction of structured information from unstructured text. A typical example is the identification of predefined relations in arbitrary text, such as marriages or company mergers. The text "The wealthy John Doe married the gorgeous Jane Doe in the beautiful countryside of Tuscany the 4th of July 2016" could be reduced to the following representation

  Marriage (John Doe, Jane Doe, 4/7/2016)

  if the relation of interest is of the type

  Marriage (*husband*, *bride*, *date*)
- **Topic Tracking**
  The goal of these algorithms is to determine whether a text could be of interest for a particular user, based on criteria such as the user's profile, his previous readings and selected keywords.
- **Summarization**
  A summarization algorithm reduces a lengthy text to few sentences that capture the essence of the whole document.
- **Categorization**
  Categorization allows to assign documents to one or more categories belonging to a predefined set.
- **Clustering**
  Clustering algorithms identify groups of similar documents (clusters) and assign them labels. In this case the labels are not predefined like the categories of categorization algorithms, but are generated on the fly based on the treated topics.

As regards the "quantitative" approach, an interesting example is described used in text tagging and annotation to analyse documents in order to identify and tag terms that correspond to domain-specific entities (for example, proper noun and numerical expressions) and then feed those terms into a classic star schema that can be queried to provide aggregated and detailed indexes.

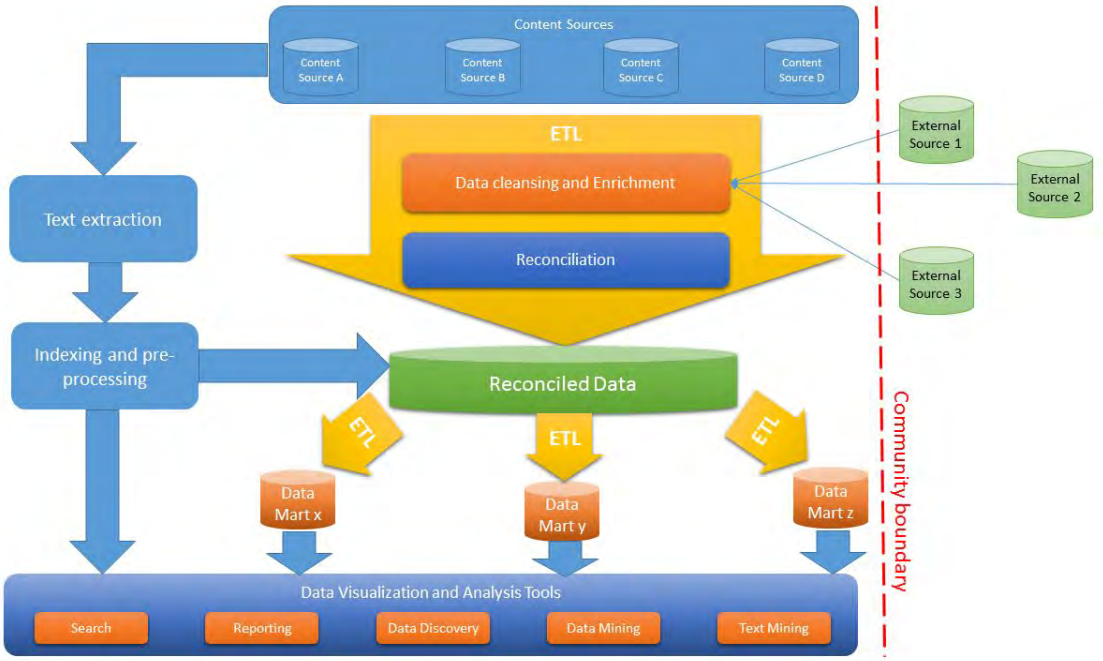


Fig. 1: data flow

In fig 1 we illustrate the data flow. At the top we have the content sources that in our case, as previously stated, are a set of repositories belonging to one or more organizations. On the left we have the flow of textual contents that are indexed and pre-processed in order to extract structured data that can be partly fed into the reconciled data base, which represents the starting point for quantitative analysis. At the centre of the figure we have the ETL procedures that cleanse, transform, integrate and enrich metadata that are then fed into the reconciled data base. Additional ETL procedures populate ad hoc data marts. The external sources on the right are obviously used to enrich metadata. Textual contents and metadata can be searched and analysed by means of the tools that are represented at the bottom pf the figure.

A Business Intelligence framework for metadata analysis

As regards source repositories, we assume that they can be harvested via OAI-PMH. **Fedora**, for example, is a well-established software for Digital Preservation repositories featuring an OAI-PMH provider. Metadata are exported in Dublin core format. Any other process for extracting metadata and contents is acceptable, although open standards are the most reliable choice in the long run.

In our technological framework, the software platform **Talend** is used to implement and execute ETL procedures that perform cleansing, enrichment and reconciliation.
The "Reconciled data" data base and the thematic data marts, are managed via PostgreSQL that, in our opinion, can be considered the most mature enterprise-level open source RDBMS.

As regards textual contents processing we think **ElasticStack** is a reliable integrated solution that features the following components:

- **Elasticsearch**, a distributed, JSON-based search and analytics engine,
- **Beats**, which acquires and sends data to Logstash and ElasticSearch,
- **Logstash**, which allows to compose the data collection pipeline,
- **Kibana**, that enables data presentation and provides a user interface for ElasticStack configuration.

The latter component can be seen as part of the "Data Visualization and Analysis Tools" box in Fig.1, which also includes Jasper (tables and charts) and Knime (Data and Text Mining).

Data discovery is an interesting new trend in data analysis and visualization, that allows data browsing and navigation through non-predefined paths. It leverages in-memory technologies and associative data bases for data storage, thus providing very low response times. To the best of our knowledge, there is currently no open source solution in this field. Commercial platforms are Tableau, Qlik, TIBCO Spotfire, Microsoft Power BI, MicroStrategy, SAP (Lumira), Platfora, Datameer, ClearStory Data, AnswerRocket, and Datawatch.

Analysis types, possible insights

A key point for any system aiming at extracting value form metadata and digital contents is to ensure that metadata quality in terms of completeness and accuracy. This can be better achieved if the quality control is performed at the content source level and if all the useful information is gathered at ingestion time.

Let's take the case of Board of Directors minutes. Much insight can be gained from text analysis and text mining, but the task becomes easier if the following information is kept as structured metadata:

- discussion points,
- decisions by type,
- yearly budgets approved,
- budget variation amounts.

In this case it is trivial to identify, for instance, a correlation between BoD decisions and organization's performance or to evaluate the frequencies of the different discussion topics. Obviously, metadata value should be, as far as possible, numeric or selected from pre-defined lists.

Best practices and European Project Examples

The launch of UNESCO programme 'Information for All'[1] provides a platform for discussing actions on information policies and the safeguarding of recorded knowledge, in coordination with the 'Memory of the World Programme', which aims to ensure preservation and universal accessibility of the world's documentary heritage.

To achieve those goals it's necessary to leverage long-term preservation and big data technologies because they can provide tools to process high volumes of data coming from different sources and represented in different formats.

Other important technological assets for the access, study and protection of cultural heritage products are: new visualization techniques, implementation of 3D models of cultural heritage, search tools for digital libraries, new approaches to digital curation and preservation.

In Europe, beside large-scale aggregators of digital collections like Europeana[2], outstanding projects about e-Infrastructures for digital cultural heritage are launched and managed by the paneuropean Network DARIAH-EU[3] - Digital Research Infrastructure for the Arts and Humanities - (18 countries, Italy included with CNR as National Coordinator) and CLARIN[4] - European research Infrastructure for Language Resource Technology.

With regard to DARIAH, which takes part in ERIC - European Research Infrastructure Consortium - and ESFRI - European Strategy Forum on Research Infrastructures - , remarkable benefits assured to digital cultural heritage communities and DARIAH-EU affiliated Projects are:

- visibility for National Research in Humanities and possibility of sharing data in a wider community; Technical environment and cooperation (e.g. virtual machines, long--term archiving, collaboration spaces etc.);
- expertise in data modeling and standards for metadata interoperability and virtual research;
- Sustainability: Research data, experiences, outcomes and publications, creating a European legal entity.

In this way DARIAH-EU Project applies for a Network with:

- close interface with Research community, their methods and questions;
- a forum to discuss current work and possible advancements, having a direct feedback from researchers expressing specific needs about digital data management and tools;
- more opportunities for financing of national and international projects.

Other interesting examples of European best practices in Cultural heritage data management and dissemination are DANS **(**Data Archiving and Networked Services)[5], and ADS (Archeology Data Service)[6], funded by the University of York (UK), which is member of Europeana Network and associated member of Digital Preservation Coalition[7]. In ADS the digital archiving of research data is entrusted to the service. ADS uses the Open Archival Information System (OAIS) reference model, which is integrated with various internal policies and procedures aimed to ensure that the data are correctly managed.

ADS main mission is "research, learning and teaching with freely available, high quality and dependable digital resources". Always keeping a watchful eye on the target audience, ADS promotes updated good practices and technical advices in managing archeological data: a correct way also to enhance ICT research on preservation and exploitation of cultural heritage.

[1] http://www.unesco.org/new/en/communication-and-information/intergovernmental-programmes/information-for-all-programme-ifap/
[2] http://www.europeana.eu
[3] http://www.dariah.eu/
[4] https://www.clarin.eu/
[5] https://easy.dans.knaw.nl/ui/home
[6] http://archaeologydataservice.ac.uk/about
[7] http://www.dpconline.org/

Finally, a short note about Vocational Education and Training: it is desireable and suggested, both by academic-scientific community and by professionals, to integrate the "big data" topic in cultural heritage and social sciences curricula studiorum, especially for the sub-fields of information science, librarianship and archival sciences.

Conclusions

In this paper we have highlighted how some types of Grey Literature assets could be leveraged to gain useful insights regarding the lives of organizations and countries. The legal obligation for long-term digital preservation currently enforced in many countries represent a great chance for building data warehouse infrastructures that collect resources and metadata from DP repositories, cleanse an enrich them and allow different types of analyses performed on ad hoc populated data marts.

References

[1] Consultative Committee for Space Data Systems Secretariat 2012, Reference model for an open archival information system (OAIS): Recommended practice (CCSDS 650.0-M-2: Magenta Book), CCSDS, Washington, DC. http://public.ccsds.org/publications/archive/650x0m2.pdf (last accessed 18/11/16)

[2] Salza, S., Il modello OAIS, Work supported by the European Community under the Information Society Technologies (IST) program of the 7th FP for RTDproject APARSEN, ref. 269977

[3] Lavoie, B., 2014 The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition), DPC Technology Watch Report

[4] Fan, W., Wallace,L., Rich,S., and Zhang, Z., 2006 Tapping the power of text mining, proceeding of communications of the ACM, pp.78-82.

[5] Parenteau, J., Sallam, R. L., Howson, C., 2015 The Rise of Data Discovery Has Set the Stage for a Major Strategic Shift in the BI and Analytics Platform Market, Gartner report ID: G00277789

[6] Prasad, K. S. N., Ramakrishna, S., 2010 Text Analytics to Data Warehousing, International Journal on Computer Science and Engineering Vol. 02, No. 06, 2010, 2201-2207

[7] Manning, C. D., Raghavan, P., Schütze, H., 2009 An introduction to information retrieval, Cambridge University Press

[8] Motta, G., Puccinelli, R., Reggiani, L., Saccone, M., 2015 Extracting value from grey literature: processes and technologies for aggregating and analysing the hidden "big data" treasure of organizations, Proceedings Grey Literature Conference 17

[9] L.Börjesson, Grey literature – grey sources? Nuancing the view on professional documentation: The case of Swedish archaeology, Journal of Documentation, vol 71(2015), nr 6, pp.1158 – 1182

[10] Guidelines on Cultural Heritage. Technical tools for for heritage conservation and management, September 2012, Conseil de L'Europe, JP - EU/CoE Support to the Promotion of Cultural Diversity (PCDK)

[11] EU grey literature. Long-term preservation, access, and discovery, Luxembourg, Publications Office of the European Union, 2012 (Cedefop Working papers, nr15), ISBN 978-92-896-1132-9, ISSN 1831-2403

[12] N. Fajrin Ariyani, U. Laili Yuhana, Generating cultural heritage metadata as linked open data in "2015 International Conference on Information Technology System and Innovation (ICITSI)", Bandung – Bali, November 16 – 19, 2015, ISBN 978-1-4673-6664-9

List of Participating Organizations

African Studies Centre Leiden, ASC	Netherlands
Alberta Health Services	Canada
American Geosciences Institute	United States
Austrian Academic Library Consortium	Austria
Biblioteca Centrale, G. Marconi; CNR	Italy
Center for Health Innovation; New York Academy of Medicine	United States
Center for Population Health Research; Lankenau Institute for Medical Research	United States
Centre National de Recherché Scientifique, CNRS	France
College of Physicians and Surgeons of Ontario, CPSO	Canada
Court of Appeals for the Second Circuit	United States
Data Archiving and Networked Services, DANS-KNAW	Netherlands
discoverygarden inc.	Canada
EBSCO	United States
East Carolina University; Health Sciences Library	United States
Federal Judicial Center, FJC	United States
Federal Library Information Network, FEDLINK	United States
Georgetown University	United States
GERiiCO laboratory	France
German National Library of Science and Technology, TIB	Germany
Grey Literature Network Service, GreyNet International	Netherlands
Indiana University, School of Informatics and Computing, IUPUI	United States
Institut de l'Information Scientifique et Technique, Inist-CNRS	France
Institute of Computational Linguistics, ILC-CNR	Italy
Institute of Information Science and Technologies, ISTI-CNR	Italy
Irvine Valley College	United States
Japan Atomic Energy Agency, JAEA	Japan
Korea Institute of Science & Technology Information, KISTI	Korea
Library of Congress, LoC	United States
McMaster University, Health Sciences Library	Canada
Metropolitan New York Library Council, METRO	United States
National Library of Technology, NTK	Czech Republic
Network and Information System Office, CNR	Italy
New York Academy of Medicine	United States
New York University; School of Medicine	United States
North Carolina State University Libraries, NCSU	United States
Nuclear Information Section; International Atomic Energy Agency, NIS-IAEA	Austria
Pratt Institute, School of Information	United States
PricewaterhouseCoopers, PwC	Netherlands
Rochester Institute of Technology, RIT	United States
Saint Petersburg State University	Russia
Slovak Centre of Scientific and Technical Information, CVTISR	Slovakia
Smithsonian Libraries	United States
TextRelease, Program and Conference Bureau	Netherlands
United Nations Dag Hammarskjöld Library	United States
University of Arizona, College of Medicine Phoenix	United States
University of Calgary	Canada
University of California, Irvine Libraries, UCI	United States
University of Florida Libraries	United States
University of Illinois at Urbana-Champaign	USA
University of Lille 3	France
University of Toronto, Institute of Health Policy, Management and Evaluation	Canada
University of Wisconsin, Milwaukee	United States
Waseda University	Japan
Wolters Kluwer, Health Learning, Research & Practice	United States
WorldWideScience Alliance	United States

Nineteenth International Conference on Grey Literature

Public Awareness and Access to Grey Literature

National Research Council of Italy

Piazzale Aldo Moro 7, Rome, Italy ▪ October 23-24, 2017

Conference Announcement

Two of the most formidable problems that have faced information through the years are its overload on the one hand and its loss on the other. These are seen as interconnected with the supply and demand sides of grey literature.

A quarter century ago, the Grey Literature Network Service joined by research communities in library and information, physics, karst and marine sciences, bio-medicine, nuclear energy, archeology, and many other scientific and technical fields set out to address this loss and overload of information.

In 1992, when the call for papers went out for the first conference in the GL-Series, the response was predominantly focused on the demand side of grey literature – that which was difficult to find and even more to access. The emphasis then lie in stemming the loss of grey literature. However, the outcome of that first conference also called attention to the equally important need for further research into the supply side of grey literature – namely its production, publication, and public awareness.

GL19 seeks to demonstrate how researchers and authors in the last 25 years have made significant inroads in responding to the loss and overload of grey literature. Likewise, this conference will seek to provide new directions in achieving public awareness and access to grey literature on an ever changing information landscape. To this end, information professionals and practitioners in the sectors of government, academics, business and industry are invited to respond to this year's **Call for Papers** reflected in the conference topics below.

Conference Related Topics

- ☐ **Exposing Grey Literature to Wider Audiences**
- ☐ **Confronting Obstacles in Accessing Grey Literature**
- ☐ **Impact of Emerging Technologies on Grey Literature**
- ☐ **Innovations in Grey Literature powered by Research Data**
- ☐ **Extracting Trusted Content from Social Media**
- ☐ **Digital Preservation, the Lifeline for Grey Resources**

GreyNet Grey Literature Network Service 1992 2017

Conference Dateline 2017

● April 15	● May 5	● May 12	● May 15	● Sept. 15	● Sept. 25	● Oct. 10	● Oct. 23-24
Close, Call for Papers	Program Committee Meeting	Authors Notified	Open, Call for Posters	Close, Early Conference Registration	Close, Call for Posters	Submission Conference Papers	GL19 Conference

TextRelease
GL19 Program and Conference Bureau

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

Nineteenth International Conference on Grey Literature

Public Awareness and Access to Grey Literature

National Research Council of Italy

Piazzale Aldo Moro 7, Rome, Italy ▪ October 23-24, 2017

Call for Papers

Title of Paper:	**Conference Topic(s):**
Author Name(s):	**Phone:**
Organization(s):	**Fax:**
Postal Address:	**Email:**
Postal/Zip Code – City – Country:	**URL:**

Guidelines for Abstracts

Participants who seek to present a paper at GL19 are invited to submit an English language abstract between 350-400 words. The abstract should deal with the problem/goal, the research method/procedure, an indication of costs related to the project, as well as the anticipated results of the research. The abstract should likewise include the title of the proposed paper, conference topic(s) most suited to the paper, name(s) of the author(s), and full address information. Abstracts are the only tangible source that allows the Program Committee to guarantee the content and balance in the conference program. Every effort should be made to reflect the content of your work in the abstract submitted. Abstracts not in compliance with the guidelines may be returned to the author for revision.

Conference Related Topics
☐ **Exposing Grey Literature to Wider Audiences**
☐ **Confronting Obstacles in Accessing Grey Literature**
☐ **Impact of Emerging Technologies on Grey Literature**
☐ **Innovations in Grey Literature powered by Research Data**
☐ **Extracting Trusted Content from Social Media**
☐ **Digital Preservation, the Lifeline for Grey Resources**

GreyNet Grey Literature Network Service 1992 2017

Due Date and Format for Submission

Abstracts in MS Word must be emailed to conference@textrelease.com on or before **April 15th 2017**. The author will receive verification upon its receipt. By mid-May, shortly after the Program Committee meets, authors will be notified of their place on the conference program. This notice will be accompanied by further guidelines for submission of full text papers, accompanying research data, PowerPoint slides, and required Author Registration.

TextRelease
GL19 Program and Conference Bureau

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

Author Information

Aloia, Danielle **151**
Danielle Aloia is Special Projects Librarian at The New York Academy of Medicine. She received her MSLS from Catholic University of America, in Washington, DC, in 2005 while working on the AgeLine Database at AARP. She has over 20 years of experience in a variety of library settings, including academic, non-profit and museum. She has been involved with collecting, evaluating, and cataloging grey literature since 2006, first at AARP and then at the United States. Dept. of Transportation. For the past 4 years she has been managing the Grey Literature Report in Public Health, produced by NYAM. Email: daloia@nyam.org

Balashova, Yuliya B. **159**
Yuliya B. Balashova is a professor of journalism at Saint Petersburg State University, Russia. She is the author of scholarly works on the history of journalism and literature, pedagogy in journalism, science journalism. Taught journalism at Western European universities. As Fulbright scholar currently affiliated with Michigan State University, USA. Email: ubalash@gmail.com

Bartolini, Roberto **109, 117**
Roberto Bartolini - Expertise on design and development of compilers of finite state grammars for functional analysis (macro-textual and syntactic) of Italian texts. Expertise on design and implementation of compilers of finite state grammars for analysis of natural language texts producing not recursive syntactic constituents (chunking) with specialization for Italian and English languages. Skills on acquiring and extracting domain terminology from unstructured text. Skills on semi-automatic acquisition of ontologies from texts to support advanced document management for the dynamic creation of ontologies starting from the linguistic analysis of documents. Email: roberto.bartolini@ilc.cnr.it

Biagioni, Stefania **97, 117**
Stefania Biagioni graduated in Italian Language and Literature at the University of Pisa and specialized in Data Processing and DBMS. She is currently a member of the research staff at the Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" (ISTI), an institute of the Italian National Research Council (CNR) located in Pisa. She is head librarian of the Multidisciplinary Library of the CNR Campus in Pisa and member of the ISTI Networked Multimedia Information Systems Laboratory (NMIS). She has been the responsible of ERCIM Technical Reference Digital Library (ETRDL) and currently of the PUMA (PUblication MAnagement) & MetaPub, a service oriented and user focused infrastructure for institutional and thematic Open Access repositories looking at the DRIVER/OpenAire 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. Since 2013 she is involved on the GreyGuide Project. Email: stefania.biagioni@isti.cnr.it

Carlesi, Carlo **97**
Carlo Carlesi, graduated in Computer Science, worked since 1970 at the IEI (now ISTI) of the Italian National Research Council in Pisa. His interest are focused in many software engineering field such as: Development of data base systems, Software quality and testing, Administration and UNIX system management, Digital library systems, Network security and management. In the year 2000-2010 he was head of the "Information Technology Security Service" of the Institute. He participated in many national and international projects, the last being: Italian Project for Research in the Antartic - (South Pole project), the aim of the project was the creation of a Multidisciplinary Integrated Information System to manage and query the Set of Antarctic Data Bases; ERCIM Technical Reference Digital Library - (ETRDL Project), The Digital Library service allows public access through Internet to the technical reports and other not published document produced by several organizations. It is currently a Research Associate of the Institute ISTI and he is involved in the following projects: PUMA - Publication Management. The Digital Library service allows public access (when permitted) through Internet to the published documents produced by CNR Organizations. GreyGuide - Guide to Good Practices and Resources in Grey Literature. An online forum and repository of good practice and resources in the field of grey literature. Email: carlo.carlesi@isti.cnr.it

Ćirković, Snježana **81**
Snjezana Cirkovic is Director of the Head Office of the Austrian Academic Library Consortium (Kooperation E-Medien Österreich). Her main responsibilities are organizing and coordinating the license agreements for electronic resources, centralized licensing for all consortia members, negotiating with suppliers, representing the consortium at international level. She holds a Master Degree in Library and Information Science from University of Belgrade, Faculty of Philology, with specialization in private libraries in the 18th century. She is currently writing PhD Dissertation with a specialization in European Cultural History during the 18th and 19th century (German-Serbian Cultural Relationships). From 2004 till 2015 she was working as a Senior Chief Librarian at the Institute for German Studies, Faculty of Philology, Belgrade University. She has developed and taught research courses of Information Literacy and Academic Writing for both M.A. and PhD students at the Faculty of Philology. Her main interests are electronic resources, open access, academic writing and scientific research.
Email: snjezana.cirkovic@gmail.com

Davis-Castro, Carla **71**
Carla Davis-Castro earned her Bachelor of Arts in Dramatic Art in 2008, Master of Public Administration and Master of Science in Library of Science in 2014 from the University of North Carolina at Chapel Hill. She joined the Library of Congress in 2015.
Email: carlayasmin.daviscastro@gmail.com

De Biagi, Luisa **162**

Luisa De Biagi got her Laurea Degree in Literature and Philosophy at 'La Sapienza' Univ. of Rome (Art history and Cultural heritage). With a Specialization in 'Archivist-Palaeographer' (Vatican School of Palaeography, Diplomatics and Archivistics at the Vatican Secret Archive) as well as a Specialization Degree in Archivistics, Palaeography and Diplomatics (Archivio di Stato, Rome) and a Degree from the Vatican School of Library Sciences. De Biagi further holds a Master in 'Business Publishing' (LUISS Management – Rome). She's been working for the SIGLE Network (System for Information on Grey Literature in Europe) since 2002, Since 2010 she's is responsible for the Italian National Referring Centre of Grey Literature at CNR Central Library 'G. Marconi' as representative of the European Network and Openarchive OpenGrey. She's taken part in 3 editions of GreyNet's GL Conference Series (GL5 Amsterdam, GL13 Washington DC, GL14 Rome and GL15 Bratislava). She's is also a member of the CNR Working Group for Cedefop-Refernet Project (Consortium for Professional Education and Training coordinated by ISFOL), the Committee for Legal Deposit Acquisition at CNR Central Library, and a member of the European Association of Health Information and Libraries (EAHIL). She's also responsible for the Library Functional Units 'Education and Training' and 'Cultural Activities Management', organizing didactics laboratories for students, professional training courses and teaching in professional trainings for librarians, students and users. Email: luisa.debiagi@cnr.it

Farace, Dominic **97**

Dominic Farace is Head of GreyNet International and Director of TextRelease, an independent information bureau specializing in grey literature and networked information. He holds degrees in sociology from Creighton University (BA) and the University of New Orleans (MA). His doctoral dissertation in social sciences is from the University of Utrecht, The Netherlands, where he has lived and worked since 1976. After six years heading the Department of Documentary Information at the Royal Netherlands Academy of Arts and Sciences (SWIDOC/KNAW), Farace founded GreyNet, Grey Literature Network Service in 1992. He has since been responsible for the International Conference Series on Grey Literature (1993-2013). In this capacity, he also serves as Program and Conference Director as well as managing editor of the Conference Proceedings. He is editor of The Grey Journal and provides workshops and training in the field of grey literature. Email: info@greynet.org

Frantzen, Jerry **97**

Jerry Frantzen graduated in 1999 from the Amsterdam University of Applied Sciences/Hogeschool van Amsterdam (AUAS/HvA) 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

Gelfand, Julia **32**

Julia Gelfand is the Applied Sciences, Engineering & Public Health Librarian at the University of California, Irvine Libraries where over the last 35 years she has performed many roles. She is active professionally and currently is a member of the Board of Directors of the Association of College & Research Libraries (ACRL), a division of the American Library Association, a member of the Science & Technology Section of the International Federation of Library Associations (IFLA) and is Secretary of Section T of the American Association for the Advancement of Science (AAAS). She writes and presents frequently on topics related to Scholarly Communication, Collection Management, Digital Scholarship, integration of multimedia in scientific literature, grey literature, social media, library as publisher. A graduate of Goucher College with graduate degrees from Case Western Reserve University, she is the recipient of many awards including the first GreyNet Award presented in 1999 and has been a Fulbright Fellow and a Thomas J. Watson Fellow. Email: jgelfand@uci.edu

Giannini, Silvia **117**

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

Goggi, Sara **109, 117**

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

Inagaki, Satomi **24**

Satomi Inagaki is a Librarian at the Central Library of the Japan Atomic Energy Agency (JAEA).
Email: inagaki.satomi@jaea.go.jp

Lin, Anthony **32**
Anthony Lin is the Head of Instruction and Collections at the Irvine Valley College Library. He holds a MSI from the University of Michigan-Ann Arbor, a BA in Spanish from California State University San Marcos, and a BS in Finance from San Diego State University. His interests are emerging technologies, effective bibliographic instruction, and collections management. Email: alin@ivc.edu

Lipinski, Tomas **11**
Professor Tomas A. Lipinski completed his Juris Doctor (J.D.) from Marquette University Law School, Milwaukee, Wisconsin, received the Master of Laws (LL.M.) from The John Marshall Law School, Chicago, Il., and the Ph.D. from the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Dr. Lipinski has worked in a variety of legal settings including the private, public and non-profit sectors. He is the author of numerous articles and book chapters; his monographs include, The Library's Legal Answer Book co-authored with Mary Minow (2003); The Copyright Law In The Distance Education Classroom (2005), The Complete Copyright Liability Handbook For Librarians And Educators (2006), and The Librarian's Legal Companion For Licensing Information Resources And Services (2012). Recent articles and chapters include, Click Here to Cloud: End User Issues in Cloud Computing Terms of Service Agreements, in Challenges Of Information Management Beyond The Cloud: 4th International Symposium On Information Management In A Changing World, Imcw 2013 (Revised Selected Papers.), with Kathrine Henderson, Hate Speech: Legal and Philosophical Aspects, in The Handbook Of Intellectual Freedom Concepts (2014), in 2013 with Andrea Copeland, Look before you License: The Use of Public Sharing Websites in building Patron Initiated Public Library Repositories, Preservation, Digital Technology & Culture and in 2012, Law vs. Ethics, Conflict and Contrast in Laws Affecting the Role of Libraries, Schools and other Information Intermediaries, Journal Of Information Ethics. He has been a visiting professor in summers at the University of Pretoria-School of Information Technology (Pretoria, South Africa) and at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Lipinski was the first named member of the Global Law Faculty, Faculty of Law, University of Leuven, Belgium, in 2006 where he continues to lecture annually at its Centers for Intellectual Property Rights and Interdisciplinary Center for Law and ICT. He is active in copyright education and policy-making, chairing the ACRL Copyright Discussion Group, a member of the ALA OITP Committee on Legislation Copyright Subcommittee, a member of the Copyright and Other Legal Matters Committee of IFLA and serves as an IFLA delegate to the World Intellectual Property Organization's Standing Committee on Copyright and Other Rights. In October of 2014 he returned to the University of Wisconsin to serve as Professor and Dean of its i-School, the School of Information Studies. Email: tlipinsk@uwm.edu

Monachini, Monica **109**
Monica Monachini is a Senior Researcher at CNR-ILC. Field of expertise: computational linguistics, computational lexicography, semantics, lexical semantics, language resources, ontologies, lexicon, terminologies, metadata, validation, methods for retrieving information in different areas (biology, environment, civil protection, oceanography, social media, humanities and social sciences, ...), infrastructural issues related to language resources. Active in many standardisation activities for harmonising lexical information. Involved and responsible of the Pisa team in many international projects for language engineering. Over the last years, she has published articles in the field of lexical resources and information extraction in different areas. Currently, she focused her activities on digital humanities. Member of various Scientific Committees; UNI delegate for ISO/TC37/SC4. Email: Monica.Monachini@ilc.cnr.it

Naughton, Robin **151**
Robin Naughton, PhD is the Digital Systems Manager for the New York Academy of Medicine. Prior to joining the Academy, Dr. Naughton was a Digital Consultant with LearningExpress, an EBSCO company, and Oxford University Press, English Language Teaching. She received an Institute of Museum and Library Services (IMLS) fellowship for digital librarianship, and completed her PhD in Information Science from Drexel's College of Computing and Informatics in 2012. Dr. Naughton is a user-centered researcher focused on human-computer interaction (HCI) and mental models, specifically how to design and build interactive systems that are useful, easy to use and enjoyable.
Email: rnaughton@nyam.org

Pardelli, Gabriella **109, 117**
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. Her interests and activity range from studies in grey literature and terminology, with particular regard to the Computational Linguistics and its related disciplines, to the creation of documentary resources for digital libraries in the Humanities. She has participated in many national and international projects including the recent projects:- BIBLOS: Historical, Philosophical and Philological Digital Library of the Italian National Research Council, (funded by CNR); - For digital edition of manuscripts of Ferdinand de Saussure (Research Programs of Relevant National Interest, PRIN - funded by the Ministry of Education, University and Research, MIUR). Email: gabriella.pardelli@ilc.cnr.it

Prost, Hélène **131**
Hélène Prost is information professional at the Institute of Scientific and Technical Information (CNRS) and associate member of the GERiiCO research laboratory (University of Lille 3). She is interested in empirical library and information sciences and statistical data analysis. She participates in research projects on evaluation of collections, document delivery, usage analysis, grey literature and open access, and she is author of several publications.
Email: helene.prost@inist.fr

Puccinelli, Roberto **162**
Roberto Puccinelli is currently head of Section I at CNR's "Information System Office" and he's been working for CNR since 2001. He has previously worked in the private sector as system and network engineer. As adjunct professor, he has held courses for the First University of Rome "La Sapienza" ("Operating Systems II") and for the Third University of Rome ("Programming and Computing Laboratory"). He graduated in Electronic Engineering at the University of Rome "La Sapienza" and holds a master cum laude in Enterprise Engineering from the University of Rome "Tor Vergata". In the past he has worked in several research projects in the field of Grid technologies both at the national and international level (executive manager of Work Package 11 within the DataGrid project – V Framework Programme, et al.). He's currently involved in the design and development of CNR's information system. In particular, he coordinates projects for the development of application systems and is responsible for the design and implementation of CNR's data warehouse. He is also responsible for CNR's Local Registration Authority management. He's currently involved in projects regarding the design and development of research product open archives and persistent identifier registers/resolvers. He is author of several articles in the fields of Grid technologies, Autonomic Computing, Software Engineering, Open Archives and Persistent Identifiers. Email: roberto.puccinelli@cnr.it

Rudasill, Lynne **114**
Lynne Rudasill holds the unique post of Global Studies Librarian and Associate Professor, University Library. Professor Rudasill's publications include edited books and book chapters as well a journal articles that focus on dissemination and use of non-governmental organization information and the resulting changes in scholarly communication, as well as user experience with library websites. She is a frequent participant in international conferences in library and information science. Lynne is a member of the Library Executive Committee, former Dean of the Military Education Council, Area Studies Division Chair and Chair of the Professional Committee of the IFLA Governing Board. She was recognized for her service to Association of College and Research Libraries with the Marta Lange/CQ Press Award.
Email: rudasill@illinois.edu

Savić, Dobrica **49, 97**
Dr. Dobrica Savić is Head of the Nuclear Information Section (NIS) of the IAEA. He holds a PhD degree from Middlesex University in London, an MPhil degree in Library and Information Science from Loughborough University, UK, an MA in International Relations from the University of Belgrade, Serbia, as well as a Graduate Diploma in Public Administration, Concordia University, Montreal, Canada. He has extensive experience in the management and operations of web, library, information and knowledge management, as well as records management and archives services across various United Nations Agencies, including UNV, UNESCO, World Bank, ICAO, and the IAEA. His main interests are creativity, innovation and use of information technology in library and information services. Email: d.savic@iaea.org

Schöpfel, Joachim **131**
Joachim Schöpfel is senior lecturer at 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 access, grey literature and eScience. He is a member of GreyNet and euroCRIS. He is also the Director of the National Digitization Centre for PhD Theses (ANRT) in Lille, France.
Email: joachim.schopfel@univ-lille3.fr

Smith, Plato L. **91**
Plato Smith is the Data Management Librarian at the University of Florida with experience in academic research libraries, digital libraries, and data management. He received his doctorate in the field of Information Science from the School of Information within the College of Communication and Information at Florida State University, Florida's iSchool, Summer 2014. From 2005 to 2012, he was Department Head for the FSU Libraries' Digital Library where he developed, populated, and managed digital collections in the FSU Libraries' digital content management system, DigiNole Repository, and electronic theses and dissertations (ETDs) institutional repository. Email: plato.smith@ufl.edu

Stock, Christiane **97**
Christiane Stock is the Head of the Monographs and Grey Literature service at INIST, in charge of the repositories LARA (reports), mémSIC (master's theses in information sciences) and OpenGrey. Member of the Technical Committee for the SIGLE database from 1993 to 2005, she also set up the national agency for ISRN (International Standard Report Number). She is member of the AFNOR expert group who prepared the recommended metadata scheme for French electronic theses (TEF). Email: christiane.stock@inist.fr

Vaska, Marcus **59**
Marcus Vaska is a librarian with the Knowledge Resource Service (KRS), Alberta Health Services (AHS), responsible for providing research and information support to staff affiliated with an Alberta Cancer Care research facility. A firm believer in embedded librarianship, Marcus engages himself in numerous activities, including instruction and research consultation, with numerous research teams. An advocate of the Open Access Movement, Marcus' current interests focus on showcasing and creating greater awareness of the LibGuides devoted to grey literature in post-secondary institutions across Canada. Email: mmvaska@ucalgary.ca

Vaska, Rosvita **59**
Rosvita Vaska is a Subject Specialist (retired) with the University of Calgary, responsible for Germanic, Slavic, East Asian, and Arabic Languages and Literatures, as well as Linguistics and Holocaust Studies. Recipient of the 2009 Order of the University of Calgary, Rosvita has been heavily involved in curriculum development, instruction, and research within her subjects of responsibility throughout her career. A firm believer in the importance of grey literature and the Open Access Movement, Rosvita is presently investigating the use of LibGuides as a grey literature document type in instructional pursuits.
Email: vaska@ucalgary.ca

INFORMATION SUPPORT OF SLOVAK SCIENCE

SCIENTIFIC LIBRARY AND INFORMATION SERVICES

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

SUPPORT IN MANAGEMENT AND EVALUATION OF SCIENCE

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

SUPPORT OF TECHNOLOGY TRANSFER

- Technology Transfer Centre at SCSTI
- PATLIB centre

POPULARISATION OF SCIENCE AND TECHNOLOGY

- National Centre for Popularisation of Science and Technology in Society

IMPLEMENTATION OF PROJECTS

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

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

Index to Authors

A-B

Aloia, Danielle 151
Balashova, Yuliya B. 159
Bartolini, Roberto 109, 117
Biagioni, Stefania 97, 117

C-D

Carlesi, Carlo 97
Ćirković, Snježana 81
Copeland, Andrea 11
Davis-Castro, Carla 71
De Biagi, Luisa 162

E-F

Ebisawa, Naomi 24
Farace, Dominic 97
Frantzen, Jerry 97

G

Gelfand, Julia 32
Giannini, Silvia 117
Goggi, Sara 109, 117
Gonda, Mayuki 24
Gruttemeier, Herbert 97

H-I

Hayakawa, Misa 24
Inagaki, Satomi 24
Itabashi, Keizo 24

J-K-L

Jones, Kyle 11
Lin, Anthony 32
Lipinski, Tomas 11

M-N

McIntyre, Lauren 91
Monachini, Monica 109
Naughton, Robin 151
Nozawa, Takashi 24

P

Pardelli, Gabriella 109, 117
Prost, Hélène 131
Puccinelli, Roberto 162

R

Rudasill, Lynne 114
Russo, Irene 109

S

Savić, Dobrica 49, 97
Schöpfel, Joachim 131
Smith, Plato L. 91
Stock, Christiane 97

V

Vaska, Marcus 59
Vaska, Rosvita 59

Forthcoming
February 2017

'Leveraging Diversity in Grey literature'

The New York Academy of Medicine - November 28-29, 2016

Publication Order Form

EIGHTEENTH INTERNATIONAL CONFERENCE ON GREY LITERATURE

Publication(s):	*No. of Copies*	*x*	*Amount in Euros*	**Subtotal**
GL18 CONFERENCE PROCEEDINGS – Printed Edition ISBN 978-90-77484-30-2 ISSN 1386-2316 ***Postage and Handling excluded****		**x**	**109.00 =**	€
GL18 CONFERENCE PROCEEDINGS - PDF Edition ISBN 978-90-77484-30-2 ISSN 1386-2316 ***Forwarded via email***		**x**	**109.00 =**	€
GL18 Conference Proceedings – Online Edition ISBN 978-90-77484-30-2 ISSN 2211-7199 ***Password Protected Access***		**x**	**109.00 =**	€

POSTAGE AND HANDLING PER PRINTED COPY*

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

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

❑ **Direct transfer to TextRelease, Rabobank Amsterdam**

BIC: RABONL2U **IBAN:** NL70 RABO 0313 5853 42, with reference to "GL18 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 WILL BE AUTHORIZED VIA OGONE/INGENICO DESIGNATED PAYMENT SERVICES

TextRelease
www.textrelease.com

GL18 Program and Conference Bureau
Javastraat 194-HS, 1095 CP Amsterdam, Netherlands
T/F +31-(0) 20-331.2420 Email: info@textrelease.com