

An International Journal on Grey Literature



Spring 2012 – TGJ Volume 8, Number 1

‘SOCIAL NETWORKING AND GREY LITERATURE’

GreyNet

The Grey Journal

An International Journal on Grey Literature

COLOPHON

Journal Editor:

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

Associate Editors:

Julia Gelfand
University of California, Irvine
UCI, United States

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

Dr. Joachim Schöpfel
Université Charles de Gaulle Lille 3
France

Gretta E. Siegel
Portland State University
PSU, United States

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

Technical Editor:

Jerry Frantzen, TextRelease

CIP

The Grey Journal (TGJ): An international journal on grey literature / D.J. Farace (Journal Editor); J. Frantzen (Technical Editor) ; GreyNet, Grey Literature Network Service. - Amsterdam: TextRelease, Volume 8, Number 1, Spring 2012. - EBSCO Publishing, FLICC-FEDLINK, INIST-CNRS, JST, NTK, and NYAM are corporate authors and associate members of GreyNet International. This serial publication appears three times a year - in spring, summer, and autumn. Each issue is thematic and deals with one or more related topics in the field of grey literature. The Grey Journal appears both in print and electronic formats.
ISSN 1574-1796 (Print)
ISSN 1574-180X (E-Print/PDF)
ISSN 1574-9320 (CD-Rom)

Subscription Rates:

€125 individual, including postage & handling
€240 institutional, including postage & handling

Contact Address:

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

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

About TGJ

The Grey Journal is a flagship journal for the international grey literature community. It crosses continents, disciplines, and sectors both public and private. The Grey Journal not only deals with the topic of grey literature but is itself a document type classified as grey literature. It is akin to other grey serial publications, such as conference proceedings, reports, working papers, etc.



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

About GreyNet

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

Full-Text License Agreement

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

© 2012 TextRelease

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

Contents

'SOCIAL NETWORKING AND GREY LITERATURE'

Science-Forums.net: A Platform for Scientific Sharing and Collaboration	5
<i>Marilyn J. Davis, Edrick Coppock, and Lance Vowell (United States)</i>	
Social Networking: Product or Process and What Shade of Grey?	14
<i>Julia Gelfand and Anthony Lin (United States)</i>	
Knowledge Communities in Grey	27
<i>Claudia Marzi (Italy)</i>	
Using Social Media to Create Virtual Interest Groups in Hospital Libraries	35
<i>Yongtao Lin and Kathryn M.E. Ranjit (Canada)</i>	
Information-Seeking Behaviour of Slovenian Researchers: Implications for information services	43
<i>Polona Vilar, Primož Južnič, and Tomaz Bartol (Slovenia)</i>	
Blog Posts and Tweets: The Next Frontier for Grey Literature	54
<i>Marcus Banks (United States)</i>	

Colophon	2
Editor's Note	4
Advertisements	
New York Academy of Medicine	34
GL14 Conference Announcement	60
GL14 Call for Papers: Tracking Innovation through Grey Literature	61
About the Authors	62
Notes for Contributors	63

EDITOR'S NOTE

GreyNet 1992-2012

Twenty-two marks the 20th Anniversary since the founding of the Grey Literature Network Service. Over the past two decades, GreyNet together with its associate members have been responsible for the program and planning of fourteen international conferences on grey literature. Research results produced by well over 300 contributing authors from more than twenty countries worldwide have been captured and preserved. These content contributions have been further compiled, edited, and made available in a host of information products ranging from conference proceedings to journal articles with further coverage in websites, newsletters, and an ever increasing range of other social media.

During these past years, GreyNet has taken the initiative in a number of research projects in order to provide stakeholders in the international grey literature community with analyses, reviews, enhanced metadata, survey results, and courseware. GreyNet has further sought to encourage and honor outstanding achievements in the field of grey literature by way of its coveted Annual Award bestowed to date on twelve information professionals in North America and Europe. GreyNet itself is the honored recipient of the Golden Candle Award in 2000 for its unique networked service and the Victorine van Schaick Prize in 2008 for The Grey Journal geared to schools and colleges of library and information science.

In lead-up to the 14th International Conference on Grey Literature scheduled this November in Rome, GreyNet further pursues a mandate from its cohorts to initiate negotiations leading up to its transition from an independent proprietorship to an international association. Readership and journal subscribers are openly encouraged to voice their comments and recommendations in this pursuit.

Dominic Farace
journal@grey.net.org

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

Science-Forums.net

A Platform for Scientific Sharing and Collaboration *

Marilyn J. Davis, Edrick Coppock, and Lance Vowell (*United States*)

Abstract

The beta website of Science-Forums.net was developed by Information International Associates, Inc. (IIA) through the Department of Energy (DOE) Small Business Innovation Research (SBIR) program. By leveraging scientist/researcher information in databases managed by DOE's Office of Scientific and Technical Information (OSTI), Science-Forums.net will allow scientists to easily collaborate on a scientific topic of interest. In addition to potentially yielding commercial applications, the functionality of Science-Forums.net is well suited to OSTI's family of scientific and technical information (STI) websites for research and development.

The objective of the SBIR—"Interactive Peer-to-Peer (P2P) Scientific Communication in the Digital Library Environment"—was to research and identify web-based tools and other concepts that will foster online interaction and collaboration among scientists and researchers. These tools can facilitate scientific discovery and innovation.

Phase I methods included (1) survey of available technologies, (2) analysis and prioritization of technologies, (3) design and planning of a foundational tool for online interaction, (4) development and implementation of approaches to obtain feedback from researchers and scientists, (5) research into privacy and communications examples and guidelines to address government privacy and survey requirements, and (6) investigation into scientists' use and interest in social media tools for research and communication.¹

Phase II has focused on (1) development of the tool into a beta-ready site for feedback from a test group of previously identified, interested researchers, (2) refinement of the privacy policy and user terms and conditions documents, and (3) further research of bibliographic/metadata controls to create and exchange documents and other content.

Motivation/Justification for the Research

This project began with the hypothesis that web 2.0 technologies can promote information exchange, collaboration, and networking among scientists (P2P) and that they can also facilitate scientific discovery and innovation. Furthermore, there was a strong belief that scientific researchers would want to collaborate in a different manner than that provided by the typical social network (e.g., Facebook, LinkedIn, Second Life, or Twitter). The system should support a broad range of information types: technical reports, grey literature, theses, research notes and papers in progress, as well as social interaction among users, blogs and forums open to the general research community, and those open only to a select group of members.

The primary goal of the project was to research and develop a prototype tool to enable interactive P2P scientific communication in the digital library setting. This prototype tool would provide mechanisms to enable scientific and research communities, as defined by common interest, to communicate based on content provided by a digital library or another collection of information (e.g., grey literature in a specific topic). By increasing the ability for collaborators to interact, this tool will improve the diffusion of scientific knowledge.

These digital libraries, or literature collections, provide a unique opportunity to facilitate communication between scientific communities of common interest. These groups can add functionality to their electronic dissemination products (i.e., web-based database search products) to allow users to communicate and collaborate using the technical information they seek as the catalyst for discussion and collaboration. To date, agencies within the

* First published in the GL13 Conference Proceedings, February 2012.

governmental digital library environment have not deployed this type of collaboration technology. This P2P communication is a less formal scientific mechanism than a technical report. Because authors typically have some mechanism for users to contact them directly, implementation of P2P communications should be straightforward.

Discussion and collaboration invariably lead to creation of new content. Collaborative technology enables informal content creation through comments and discussion, along with the more formal creation of published and grey literature, using document and bibliographic/metadata formatting interchange standards. Mechanisms to perform a literature search and to access, import, share, view, and cite other published and grey literature are also necessary to facilitate creation of new content.

Other Benefits

Ila anticipates that the project will result in a P2P system that is easy to use, configure, and administrate. The approaches developed and implemented for this project are important for government agencies and for corporations providing public information to various consumers on widely accessible terms, as well as through more restricted access. Any research institution or information provider that makes a body of STI available will greatly benefit from this project. This communication mechanism will enable members of scientific communities to discover others with common interests, facilitating a broader perspective on specific topics, and increasing and enhancing the amount and quality of information available for a specific community of interest.

This system will also accelerate knowledge discovery. Having the ability to communicate instantly with peers and to identify a broader field of members within communities of interest will facilitate discovery and synthesis of STI. Furthermore, enabling the development of new scientific grey literature in standard bibliographic/metadata formats will facilitate retrieval by librarians and other information professionals. Not only is this concept useful to traditional government research agencies, but it is also applicable to universities, pharmaceutical companies, and all organizations involved in scientific research.

PHASE I

Experience, Goals and Accomplishments

Based on our analysis of the requirements during Phase I, we identified appropriate DOE libraries and other agency libraries to participate in the prototype. We conducted joint application design (JAD) sessions to determine the type of P2P communications desirable in this environment. We used a structured questionnaire to poll representatives from selected agencies and national laboratories, as well as local information analysts, to determine how they currently collaborate and how they would like to collaborate in the future. Ila developed and refined requirements through an iterative effort with DOE and other agency stakeholders. Phase I also included a comprehensive survey of available commercial technologies and standards to determine what commercial-off-the-shelf (COTS) or shareware products provide P2P communications to satisfy some or all requirements. Technologies in development at universities, within DOE, and within other government agencies were also surveyed. Ila anticipates that available technologies will provide 60 percent of the solution.

User Input Process - Interviews, Polls, etc.

A two-pronged approach was used to determine baseline requirements.

1. We interviewed individuals in a small, hand-selected focus group in person to define a useful set of collaboration tools. In one of the focus group interviews, we asked a DOE nuclear researcher his opinion of new social media. A few of his quotes follow:
 - "If there are ways to have a social networking site devoted to narrow topics, that might be of value. So, making that easy for someone to establish something like that. But it would have to be bottom-up, researcher-driven."0
 - "I'm working now in the area of thermodynamic modeling of nuclear fuel. I have some colleagues in that area. I hear from them occasionally. It would be nice to foster a little more communication on that."

- “There is a potential that we’re overlapping in work, and we want to avoid that. So-and-so has developed an understanding of this chemical system. I don’t want to do the same thing. I want to leverage what he’s already done, and vice versa.”
- 2. We identified a larger group to answer a questionnaire and provide feedback on a regular basis. To identify this group, emails were sent to more than 11,000 DOE researchers and authors explaining the project and asking for their assistance. Based on their responses, 136 respondents were selected to participate in the feedback.

Summary

There were three objectives in Phase I: (1) determine the high-level requirements for the system, (2), conduct a survey of COTS/open source systems, and (3) produce a prototype for evaluation. These goals were met, with results as discussed below.

Requirements

Several critical themes were almost universal among the respondents. Simplicity and ease of use were the first requirements identified by interviewees. In virtually all instances, the researchers said they did not have the time or the inclination to wade through a system of complex screens and infrastructure to try and use the system. They wanted to keep the focus on their research and publication rather than become the master of a complex tool. Another common theme was the need to efficiently access, import, and share content (especially from library collections), in-process work, lab notes, and other unpublished information. While this was not a focus in the initial phases of the project, it quickly became apparent that this system could be an extremely useful tool for the GreyNet community by providing the platform to make collections of grey literature available to select communities of interest using the uploaded information and other databases as a basis for discussion and collaboration. Likewise, forum members can upload their own collections of grey literature to use as a basis for discussion in their forum.

Sensitivity to the conflict “between collaboration and competition inherent to science”² was another theme. Interviewees want to be confident that adequate security safeguards are in place to ensure that information is shared only by trusted individuals who are members of their forum.

Market Survey

When the initial marketplace review was conducted early in Phase I, very few scientific collaboration products were available. Almost all of products were narrowly focused on specific scientific communities. During Phase I, extensive web research was conducted, and university and national laboratory work in this area was examined. While some products were partially related, in the end, no technologies were uncovered to fully solve the problem.

Prototype

The other primary purpose of the research was to develop a limited prototype with some of the planned functions for evaluation by DOE. The prototype was to demonstrate the concept’s usefulness and desirability. The prototype was constructed and shared with DOE, who concurred that the concept was viable and would be a valuable addition to existing OSTI STI web products.

Our Phase I final report states: “The project team designed, prototyped, and integrated a web-based application feature called ‘Author and Subject Clustering’ into the OSTI Information Bridge web site. Author and subject clustering allows researchers to view a list of prominent authors and subjects for each topical search they perform.”³

PHASE II

Prototype Development

For the enhanced prototype in Phase II, we developed and refined requirements through an iterative process among DOE stakeholders and Ila. Therefore, the initial requirements were expanded as needed by analyzing, documenting, and reaching consensus on requirements through prototype functions in the following areas:

- User interface
- Model to identify common interest groups
- Cost-effective, real-time, interactive P2P communication
- Capability to access, import, create, search various documents/content
- Privacy
- Security
- Data protection
- Reporting

The main focus of Phase I was to develop a questionnaire for use throughout the user community to determine the system's operational requirements. Based on this feedback, a system was designed to enhance collaboration within the scientific community, and a limited, prototype was developed to highlight authors within DOE online repositories.

For Phase II, we revisited a survey of currently available commercial technologies, including COTS or shareware products with P2P capabilities that may be adapted to the digital library environment. We expected and confirmed that there was considerable development in P2P collaboration between Phase I and Phase II surveys. In fact, some commercial products were released during this time. However, they were either so complex that target scientists were unlikely to adopt them, or they did not include important functionality needed to ensure effective performance for this audience. We did find that a significant number of shareware or open source products were highly customizable out of the box to fulfill many project requirements.

Before selecting a technology for the enhanced prototype, we identified barriers to adoption for the scientific community. Along with thorough evaluation of comments from the Phase I interviews and questionnaires, we also analyzed studies, literature, and online discussions of scientists. A recurring theme is summarized in this quote: "Most scientists were reluctant to invest more than a very small amount of time to learn to use new technologies unless the benefits were substantial and related directly to their research." This observation is consistent with earlier conclusions, underlining the importance of creating an intuitive, simple prototype with valuable tools and functionality. It also reinforces the conclusion that the flexibility of today's technology—doing everything for everybody—results in complex tools that are difficult for many scientists to use.⁴ This exercise resulted in a list of obstacles that the enhanced prototype would need to overcome to ensure success (see Appendix 1).

Technology

During the Phase I technology survey and review of the existing technologies, we determined that there was no COTS product to meet project needs and that a prototype was needed. Therefore, we had to determine whether to develop something from the ground up, or whether we should select an open source framework to offer some initial desired functionality and to provide a development framework so that we could quickly add and evaluate new features. After a thorough evaluation of existing technologies, we focused on an open source content management system (CMS) to provide the most functionality out of the box while also providing a development environment in which features can be added as needed. We considered Ning, Drupal, Mambo, Joomla!, Taverna, and SupportMaster. Drupal was chosen based on several factors. Drupal is open source software maintained and developed by a community of hundreds of thousands of users and developers. It is distributed under the terms of the GNU General Public License (GPL), so it is free to download and share, and users can also contribute back to the project. This open development model ensures that people are constantly working to make sure Drupal is a cutting-edge platform that supports the latest technologies that the web offers. In addition, OSTI had selected Drupal for several applications

(e.g., OSTIblog), so local expertise is available on the product. The Drupal CMS was chosen to allow for rapid prototyping of basic functions. Once it was determined that the tool met expectations for rapid prototyping, modifications and add-ons were made to make Drupal accommodate unique functional requirements and address performance issues. Drupal also paid big development dividends since there are literally thousands of contributed modules available to help solve unique requirements. However, because Drupal is so rich and flexible, performance is impacted. As established earlier, poor performance is not likely to be tolerated by busy scientists. However, the intent was to use Drupal as the prototype, rewriting the system if the concept is commercially viable.

Because a major part of the system's functionality is to serve as a repository (library) for scientific research information, we determined that it was important to use open access standards for importing documents in the library function. We analyzed document metadata using digital object identifier (DOI), ISBN, Dublin Core, (along with OSTI DOE documents), and other standards for interoperability. The Library of Congress (LOC) website was used as a model for document interchange and standards. The result of the research and development is a web-based CMS that enables self-forming groups where researchers can create useful websites around scientific topics of interest to those groups. The project team named the website and system "Science-Forums.net."

Beta System Functionality

The purpose of the "Profiles" function is to encourage and support interaction among users of Science-Forums.net and ultimately of OSTI web portals. Profile information enables users to accomplish the following:

1. **Verify the identity of other users.** This is especially valuable when users have similar last names. For example, when a user's affiliation is provided in a profile, it can help end-users determine which "Smith" is the author of a particular document.
2. **Specify other names under which authors have published.** When an author has published under more than one name, the ability to make other names known can help users identify other works by the same author.
3. **Provide more detailed background.** Profiles can include information on education or work history to let other users know more about a particular user's credentials or experience with a subject. This is often of interest to researchers performing topic or author searches, viewing answers to questions, and understanding the context of online discussions.
4. **Declare research areas and fields of interest.** Profiles with this information can facilitate connections and dialog on topics of mutual interest.
5. **Broadcast a user's interest in collaborative opportunities.** This information can help users who are seeking or offering partners for proposals.
6. **Identify and locate researchers with similar interests.** This occurs in social networking websites and can be emulated here.

Trust is a critical element for effective online interaction. Without trusted communication, collaboration cannot take place. In order to achieve the project vision of researchers meeting, talking, and collaborating online, it is essential to provide a reliable level of trust for those interactions. This can be achieved through identification and verification of forum participants. The use of profiles is widely accepted and is scalable for different purposes. Profiles can be used on networked systems to identify and nominally authenticate users.

During the design phase, we considered providing a trust service that authenticates new participants by verifying some of the user-provided background information. We determined that this process would be too resource-intensive for the prototype, so the trust service will be explored by stakeholders for inclusion in subsequent increments.

Other trust services that Science-Forums.net profiles can provide include introductions and messaging between previously unacquainted users while protecting user privacy.

Researchers/users that express interest in providing feedback will be the first invited to create profiles to test the forum system.⁵ Authors with articles in OSTI collections may also be invited to create forums and profiles. When the Science-Forums.net framework is integrated with OSTI web portals, a profile link will be available for authors who have provided profiles.

We considered a broad range of options for the type of information to include in user profiles, ranging from a minimal amount to a very robust profile like that used for LinkedIn or Facebook. Based on feedback during requirements analysis, a baseline model was chosen at this stage, with added features to address unique requirements of scientists and professionals. After the user creates a user profile and initially logs onto the system, then he or she is presented with a welcome screen, which provides the user with access to the Forums screen.

At this point the user can search or browse the existing forums or create his or her forum. There are two types of forums that can be created - public and private. In a public forum, any user can join, search, view and contribute. In a private forum, membership is controlled by the forum creator and can only be viewed and contributed to by invited members. Once a forum is created and selected, the user is presented a screen in which the navigation bar can now change to display the following headings.

- **Library:** the main repository for existing technical content. The user can add content directly from DOE holdings, search and add content based on ISBN or DOI, cut/paste from existing metadata (e.g., BibTex or EndNote), and manually enter metadata for user-created documents. If they are available, the user can upload the full-text documents associated with the metadata in common document formats such as pdf, xls, doc, jpeg, etc.
- **Posts/Discussions:** blog posts, announcements, Q&A, and other communications for review, feedback and discourse. Other members can comment on posted entries, and users can view threads of conversations on a topic and respond to one or more previous posts. References or links to related information can be included in the body of the comment to support any point of view.
- **Calendar:** allows forum members to add event entries to mark meeting dates, conferences, or other events of interest.
- **Members:** allows the user to view all forum members and to manage privileges. In addition, it allows the user to invite colleagues to join the forum.

The system also supports other features, such as notifications via RSS feeds, forum administration, user feedback, searches, and user support such as FAQs and contact information. In addition, a full user profile management subsystem is provided.

CURRENT STATUS

The system is nearing the end of beta testing and is under consideration by several scientific groups for adoption as their collaboration portal.

FUTURE DIRECTIONS

Depending on the level of system adoption and the availability of future funding, several enhancements have been identified and will be included in the system. In addition, if the level of performance does not meet users' needs, we will consider a customization that relies less on Drupal modules to provide better performance.

APPENDIX 1

Barriers to Adoption of Social Networking by Scientists, Businesses, Workers

Type of Barrier	Description / Excerpt from Source
Benefits of social networking are not obvious to researchers	The major barrier to take-up of web 2.0 tools and services is lack of clarity – even among some frequent users – as to what the benefits might be. The costs of adoption are not always trivial, and unless researchers receive active support and see clear and quick benefits, they tend to keep to the tools and services that they know and trust. Moreover, the rapid development and proliferation of web 2.0 services mean that it is hard to keep track of them, or assess their potential benefits. ⁶
	Some applications promise a lot of value, but people may not see the immediate value. ⁷
	Scientists are not really interested in social networking as an end in itself. They network to boost productivity. ⁸
Must be easy to use, and take little time/effort to learn	While the technologies used to collect and analyze the data often must be cutting edge, specifically collaborative technologies are often best served by technologies that are simple, require very little learning, and are already easily accessible (e.g. wikis, and telephones). Most scientists were reluctant to invest more than very small amount of time to learn to use new technologies unless the benefits were substantial and related directly to their research. ⁹
	Widespread adoption of web 2.0 services by researchers depends on their being intuitive and easy to use , and incremental in building on existing practices. Above all, they must offer clear advantages to users, and near zero adoption costs. ¹⁰
Perceived lack of Quality and Trust (without traditional peer-review)	But a second major set of barriers revolve around perceptions of quality and trust. Both as producers and consumers of information, researchers seek assurances of quality; and many of them are discouraged from making use of new forms of scholarly communications because they do not trust what has not been subject to formal peer review. A significant minority of researchers believe that peer review in its current forms will become increasingly unsustainable over the next five years, and nearly half (47%) expect that it will be complemented by citation and usage statistics, and user ratings and comments. But at present they do not see such measures as an adequate substitute for peer review. Trust is also a concern for researchers who are producing, rather than consuming, information; they are cautious about sharing results and findings in a medium which, as yet, has no standardised way to formally attribute authorship. ¹¹
	But the major disincentive for many researchers may be lack of trust. Both as creators and consumers of content and services, researchers seek assurances of quality. Our study indicates that many researchers are discouraged from using new forms of scholarly communications because they do not trust what has not been subject to formal peer review. These findings are consistent with other studies (e.g. Ware and Monkman 2008) which suggest that researchers seek assurances of quality above all through peer review, and that they do not see citation counts, usage statistics or reader ratings or other 'wisdom of the crowds' tools as providing an adequate substitute. ¹²
	David said that some of the barriers to adoption include the lack of reputation information, ¹³ lack of confidence in individual networks, ¹⁴ and concerns that personal data might be sold, ¹⁵
	Users may also be suspicious of a site's commercial intentions. ¹⁶
	Data used in networks need provenance and structure - some kind of hosting institution or scientific society verifying the credibility of data. ¹⁷
	Sites with no checks on data run the risk of offering less valuable information. ¹⁸
Waste of time	some researchers regard blogs, wikis and other novel forms of communication as a waste of time or even dangerous ¹⁹
	Some non-users go further, and believe that novel forms of scholarly communications bring no benefits or are even a 'waste of time.' 'I'd rather spend the time thinking about what I'm going to do next rather than spend it telling others what I'm doing... I think it's definitely a younger person's thing.' ²⁰
Needs critical mass of users to be effective	These problems are exacerbated by the fragmentation of the user-base: few services have yet achieved the critical mass needed to achieve the positive network effects that stimulate pervasive use by particular communities. Researchers may well be right to defer a decision to take up a particular service until they are sure that large numbers of their colleagues have done so. ²¹
	But there is some debate about whether many of the web 2.0 services for researchers – particularly social network services – provide sufficient added value to stimulate widespread adoption. ²²

Type of Barrier	Description / Excerpt from Source
	Moreover, the plurality of services results in fragmentation of the potential user base, which is especially problematic when benefits are closely related to number of users. Researchers may well defer a decision to adopt until they are sure that large numbers of their colleagues have done so. Thus the advantages for late movers may outweigh those for early adopters. ²³
	Social networking does not have value if you are just talking to yourself. (Resseque, 2010) ²⁴
	no one site provides tools or features valuable enough to lure a majority of busy scientists ²⁵
	What can they offer that more established sites, such as Facebook, don't? ²⁶
Applications not integrated	David said that they've tried Yammer here and a wiki there, but when the applications are not combined in one product, it splits people's attention and the usage falls off. ... Some don't know how to use applications together. ²⁷
Change may conflict with existing practices	Changes are usually incremental, for a number of reasons. Introducing changes without input from users may conflict with existing practices and be rejected. Moreover, it takes time for new practices to develop around new features, especially if these radically challenge existing disciplinary patterns of use. Providers therefore generally seek to align new features with existing services and patterns of usage. Thus, the fora on Nature Networks are very much in the style of conventional bulletin boards, with implicit social rules of behavior and a reasonably active moderator. ²⁸
	Adoption of web 2.0 tools and services, and of novel forms of scholarly communication associated with them, has reached only modest levels up to now. Use is both fragmented and uneven, and tends to support well-established practices. ²⁹
Institutional policies can discourage use	We also found some evidence that emerging institutional policies may act as a barrier: 'In our university we have a guideline on what may or may not be put onto the blog. I have to agree that something needs to be saved and I don't want people to say: we just discovered X.' ³⁰
Conflict between cooperation and competition	Some scientists may be wary of sharing too much, which probably hinders network adoption. There is tension between collaboration and competition inherent to science. ³¹

References

- ¹ Coppock, Edrick, O'Dell, D. (2009), Interactive Peer-to-Peer Scientific Communication in the Digital Library Environment, DOE FY2008 SBIR, DE-FGO2-08ER85096, Final Report, Department of Energy, May 28, 2009.
- ² Gewin, Virginia. (2010, 2011), Collaboration: Social networking seeks critical mass. – In: Nature, and Naturejobs.com, Dec 15, 2010, <http://www.nature.com/naturejobs/2010/101216/full/nj7326-993a.html>.
- ³ Coppock, Edrick, O'Dell, D. (2009), Interactive Peer-to-Peer Scientific Communication in the Digital Library Environment, DOE FY2008 SBIR, DE-FGO2-08ER85096, Final Report, Department of Energy, May 28, 2009.
- ⁴ Lee, C. P. and M.J. Bietz. (2009), Barriers to the adoption of new collaboration technologies for scientists. Position paper for The Changing Face of Digital Science: New Practices in Scientific Collaborations workshop, the ACM Conference on Computer-Human Interaction (CHI) – In: matthewbietz.org, Apr 5, 2009, <http://www.matthewbietz.org/blog/wp-content/uploads/chi2009-scientificcollaborationsposition.pdf>.
- ⁵ Coppock, Edrick, O'Dell, D. (2009), Interactive Peer-to-Peer Scientific Communication in the Digital Library Environment, DOE FY2008 SBIR, DE-FGO2-08ER85096, Final Report, Department of Energy, May 28, 2009.
- ⁶ Proctor, R., R. Williams, and J. Stewart. (2010), If you build it, will they come? How researchers perceive and use web2.0. London: Research Information Network.
- ⁷ Resseguie, David. (2010). Personal observations at EPSCoR monthly meeting on Exploiting the Use of Social Networking to Facilitate Collaboration in the Scientific Community, Nov 2, 2010.
- ⁸ Gewin, Virginia. (2010) "Collaboration: Social networking seeks critical mass." – In: Nature, and <http://www.nature.com/naturejobs/2010/101216/full/nj7326-993a.html>. Naturejobs.com, Dec 15, 2010.
- ⁹ Lee, C. P. and M.J. Bietz. (2009), Barriers to the adoption of new collaboration technologies for scientists. Position paper for The Changing Face of Digital Science: New Practices in Scientific Collaborations workshop, the ACM Conference on Computer-Human Interaction (CHI) – In: matthewbietz.org, Apr 5, 2009, <http://www.matthewbietz.org/blog/wp-content/uploads/chi2009-scientificcollaborationsposition.pdf>.
- ¹⁰ Proctor, R., R. Williams, and J. Stewart. (2010), If you build it, will they come? How researchers perceive and use web2.0. London: Research Information Network.
- ¹¹ Ibid.
- ¹² Ware, Mark and Monkman, Mike. (2008), Peer review in scholarly journals: Perspective of the scholarly community – an international study. Publishing Research Consortium.
- ¹³ Resseguie, David. (2010). Personal observations at EPSCoR monthly meeting on Exploiting the Use of Social Networking to Facilitate Collaboration in the Scientific Community, Nov 2, 2010.
- ¹⁴ Gewin, Virginia. (2010, 2011), Collaboration: Social networking seeks critical mass. – In: Nature, and Naturejobs.com, Dec 15, 2010, <http://www.nature.com/naturejobs/2010/101216/full/nj7326-993a.html>.
- ¹⁵ Ibid.
- ¹⁶ Ibid.
- ¹⁷ Ibid.
- ¹⁸ Ibid.
- ¹⁹ Proctor, R., R. Williams, and J. Stewart. (2010), If you build it, will they come? How researchers perceive and use web2.0. London: Research Information Network.
- ²⁰ Ibid.
- ²¹ Ibid.
- ²² Ibid.
- ²³ Ibid.
- ²⁴ Resseguie, David. (2010). Personal observations at EPSCoR monthly meeting on Exploiting the Use of Social Networking to Facilitate Collaboration in the Scientific Community, Nov 2, 2010.
- ²⁵ Gewin, Virginia. (2010, 2011), Collaboration: Social networking seeks critical mass. – In: Nature, and Naturejobs.com, Dec 15, 2010, <http://www.nature.com/naturejobs/2010/101216/full/nj7326-993a.html>.
- ²⁶ Ibid.
- ²⁷ Resseguie, David. (2010). Personal observations at EPSCoR monthly meeting on Exploiting the Use of Social Networking to Facilitate Collaboration in the Scientific Community, Nov 2, 2010.
- ²⁸ Proctor, R., R. Williams, and J. Stewart. (2010), If you build it, will they come? How researchers perceive and use web2.0. London: Research Information Network.
- ²⁹ Ibid.
- ³⁰ Ibid.
- ³¹ Gewin, Virginia. (2010, 2011), Collaboration: Social networking seeks critical mass. – In: Nature, and Naturejobs.com, Dec 15, 2010, <http://www.nature.com/naturejobs/2010/101216/full/nj7326-993a.html>.

Social Networking: Product or Process and What Shade of Grey?*

Julia Gelfand and Anthony Lin (United States)

Abstract

Social networking debuted in 1997 and is now an established and common method of communication and is increasingly related to and supportive of academic publishing, scholarship and generating new information. Some of the most mature and popular sites are Facebook, Bebo, Twitter, Linked-In and Plaxo plus many more specialized examples. As many professional societies and individuals choose to develop a presence on social networking sites (SNSs), the utility of them has become more valuable and ubiquitous. With emerging forms of technology to provide personal space and networking opportunities, the sites gain a new presence blending information products with new means to be discovered and searched. This paper explores how academic libraries are implementing a range of social networking activities to augment their online presence through traditional websites and launch new ways for their users to find, discover, access, navigate, evaluate, create and apply content. Libraries are also extending access with greater mobile optimized sites for Smartphones (iPhone, Droid, and Blackberry) and other emerging technologies that enhance information in utilizing and managing data, images, audio and streaming media. Today, library collections and services are being transformed due to electronic publishing, data curation, preservation and archiving efforts in order to allow users dependable 24/7 access to resources with potential for interactive communication. In an increasingly global and mobile society, the new social communication mediums reduce boundaries, transcend the digital divide and invite more transparency and participation by shadowing the grey parameters in the scholarly community and redefining publishing outputs and opportunities. This evolution did not happen without cautious regard to intellectual property, privacy and confidentiality. The new by-products are not always tangible but remain instructive and promote innovation within communication as a new adopted and accepted form of authorship and creative expression. By adding the social communication features of this kind of networking that promotes introduction, networking, commentary, critique, discourse, sharing, and the building of new communities, grey content is enhanced and the sources of access multiply with an open and entrepreneurial future for SNSs.

Let's begin with a story, not a fairytale but a true fact-based memory.

A not quite long time ago in a far away land, there existed a feudal system of Publishers who sought to control the largest number of Users as their loyal subjects (archeologists would later refer to this prehistoric era as Web 1.0). Users were happy for a while until one day they discovered that they too, could become Publishers, quite easily in fact, and form their own kingdoms and communities.

They created a new land where content is democratized and where each User could become their own king. New social websites and services soon enabled them to publish royal decrees (called Blogs and Podcasts), vote for content they Digg, discover old friends on Social Networks, Poke (and even SuperPoke) each other, become Internet Famous, gain Followers, and in a few cases become even larger and more powerful than the great Publisher kingdoms of yore. They named this new utopia Web 2.0 and immediately started to blog about how cliché the name was. Welcome to the Social Web.¹

* First published in the GL13 Conference Proceedings, February 2012.

A Short History of Social Media



Courtesy of Socially Aware,
the social media law update; to subscribe,
please visit www.mofo.com/sociallyaware.

MORRISON | FOERSTER

©2011 Morrison & Foerster LLP | mfo.com

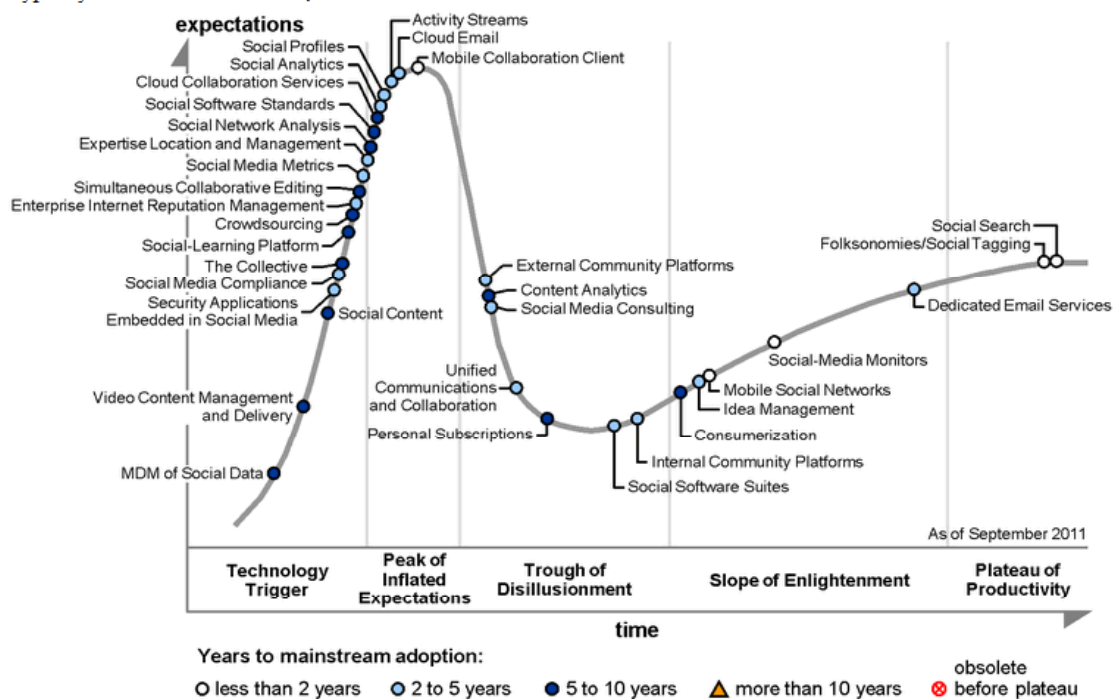
Background

The relationship between Grey Literature and Social Media is not at all unusual. Grey Literature has been known to include the previously challenging documentation, records, and communications in all formats including oral histories, news and broadcasts, correspondence, diaries, journals, reviews, other primary source content and more recently

websites, blogs, examples of photojournalism, videos, images, and other forms of commentary. The enhanced technologies which are now more ubiquitous in everyday life make for an intensive lifestyle change that allows individuals a way to present themselves, build and conduct independent relationships and offers personal empowerment that has never been so intensely realized prior to the development of the Internet. It often provides same or real time experiences and broadcasts across all geographical barriers as long as one has connectivity and anyone of many devices required to participate. Here we are at the end of 2011,

- Nearly 20 years after the birth of the World Wide Web
- 13 years after the launch of Google Search
- 8 years after the start of the first social networking site
- 6 years after the first YouTube video
- 4 years after the introduction of the first touch screen smartphone
- 3 years after the opening of the first “app” store
- And a little over a year after the first iPad sale²

Hype Cycle for Social Software, 2011



Source: Gartner, September, 2011(captured October 30, 2011)

Social media has been defined by Gartner as:

“a set of technologies targeted at forming and enabling a community of participants to productively collaborate. The ability of social-media technology to enable mass collaboration differentiates it from other collaboration technology and practices. IT tools to support collaboration have existed for decades. But social-media technologies, such as social networking, wikis and blogs, enable collaboration on a much grander scale. They enable potentially hundreds of thousands, even millions of people to simultaneously create content, share experiences and build relationships.”³

These contemporary methods of social media are indeed different from other forms of exchange in that they can be saved, traced, shared, easily accessed and have unspecified properties challenging authorship, rights, permissions, and intellectual property. As much as

they are technologies, they are tools. Suffice it to say, social media has redefined the information landscape. Print advertising revenue streams are drying up quickly due to readership preferences and patterns. The upswing of e-publishing and digital formats offers the added power of interactivity, commentary and opportunities for readers' feedback. Anyone can be an author or contributor and share their opinion with whatever bias with the entire world if they choose. The Global Web Index reported three clear trends in the consumer adoption of the Internet (not a medium):

1. Social Media has reached mass maturity – shift from massive growth and active social consumers to 'real-time' technologies – moving from creating content and publishing to sharing or from creator to distributor models for both 'traditional' media and professional content- utilizes tweets, status updates, etc
2. Open browser-based web is losing out to packaged internet platforms such as mobile apps, internet connected televisions, tablets, eReaders, gaming and video platforms
3. Continued aggressive growth of packaged platforms as traditional style content become a core part of the consumer online experience⁴

It is no surprise that as we witness the changes in how media is delivered, cable television, satellite, radio and newspapers as we currently know them are in danger of becoming obsolete. A more "packaged" Internet may transform the way we get online, the content we consume, and the ways we can create, share and communicate.⁵ We also see a convergence taking place within new social media as Twitter becomes increasingly visible and important and the academic community abandons or relies less upon RSS feeds for updates than on twitter feeds, choosing to find twitterers who keep current, participate in timely conversations/debates and provide links to sources relevant to personal interests, both individuals and relevant organizations. These methods of keeping up with technology will be personally challenging as well as a goal for the workplace to handle.⁶

Social Media as Grey Literature

Grey literature has always encompassed new ideas and captured information outputs regardless of formats and as technologies have emerged, embraced them to the extent that they can be saved, preserved and archived. In the last twenty years of the digital revolution, we have been immersed in the knowledge generation era followed by one in knowledge management, tapping into the vast potential that hypermedia and the Internet allows. Higher education and research benefited from this in a very transformational way as the structure of social construction expanded to include new communication methods. Previously, a world of standard grammar peppered with occasional colloquialisms and nuances, new lingo, and expressions would be common. Today, new languages to accommodate social media have been birthed, text messaging, for one has its own dictionaries and a lingua franca of the hand-held device has determined ways to create its own ethnography. In one of the first books that chronicled this movement of Social Media the compiled conference papers allowed the editor, Edward Barrett to coin the term, 'sociomedia,' "to signify when we design computer media we are hardwiring a mechanism for the social construction of knowledge."⁷ The papers in this early collection (1989) explore the role of the computer and what is information and he writes, "Information about a text becomes as important as the primary text itself."⁸ An aspect of daily life that has really assumed added visibility due to social media is the expansion of religious identity via the new communities established through these online venues. (Church of Facebook) As the *Church of Facebook* describes itself, the hyperconnected are a redefining community and encouraging the pursuit of authentic relationships with spiritual beings outside of the secular domain.⁹

No subject, discipline, activity, workflow is left untouched by social media and its potential. The business community's thread to survival is keenly dependent on its mastery, creative and appropriate applications. It is considered by many current writers and historians of the day, that the omission of social media will delete you and your enterprise from the public radar.

Still considered in its infancy, social media remains to be all about potential and users will have to be flexible and able to cope with frequent change, and a certain amount of ambiguity.

While we were finishing this paper, the news of the passing of two extraordinary and highly influential people who contributed to this topic of information sharing communities and networks cannot be overlooked. Steve Jobs, founder of Apple and John McCarthy, known as the father of artificial intelligence. Without their innovative spirits, our today would be a lot different.

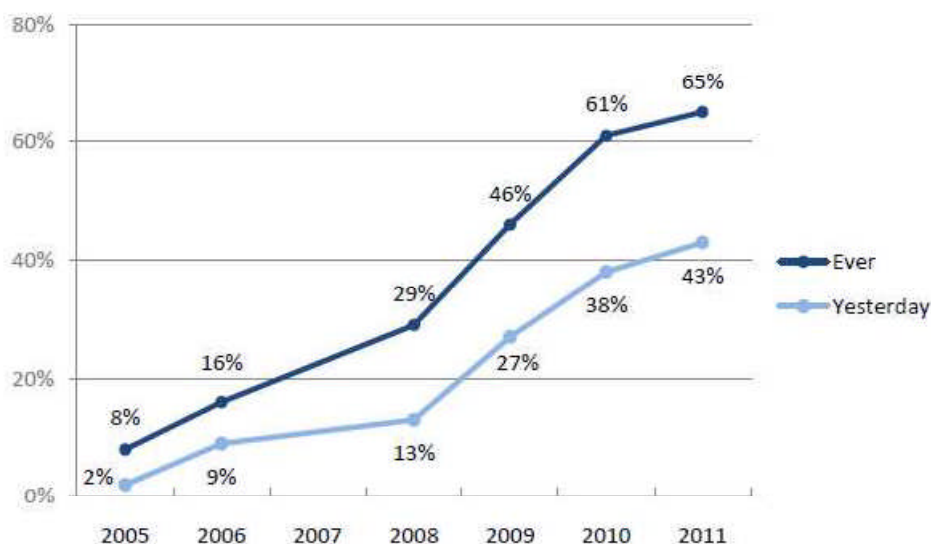
As librarians, our generation is focused on knowledge creation and we commend David Lankes for including a “thread” on this in his much acclaimed, *Atlas of Librarianship*. He quotes Melvil Dewey, “The time was when a library was very much like a museum, and a librarian was a mouser in musty books...The time is when a library is a school, and the librarian is the highest sense a teacher.”¹⁰ The Atlas is intended to have a worldview of librarianship not founded on materials but outcomes and learning with concentration on the librarian rather than the institution.¹¹ The Mission Statement, “The Mission of Librarians is to Improve Society through Facilitation Knowledge Creation in their Communities,” is a figurative illustration about how that can be done with different focus points and the one selected for this paper, emphasizes conversation and should be viewed vertically or top to bottom.¹²

How we perceive social media networking

Increasingly, people are using social media for a variety of different things. According to the Pew Internet Research Foundation, more and more users are using social media as a daily part of their lives¹³. As can be seen from the chart below, social networking site use by adults starts from a low 8% in 2005 to nearly 65% in 2011.¹⁴

Social networking site use by online adults, 2005-2011

The percentage of all adult internet users who use social networking sites since 2005



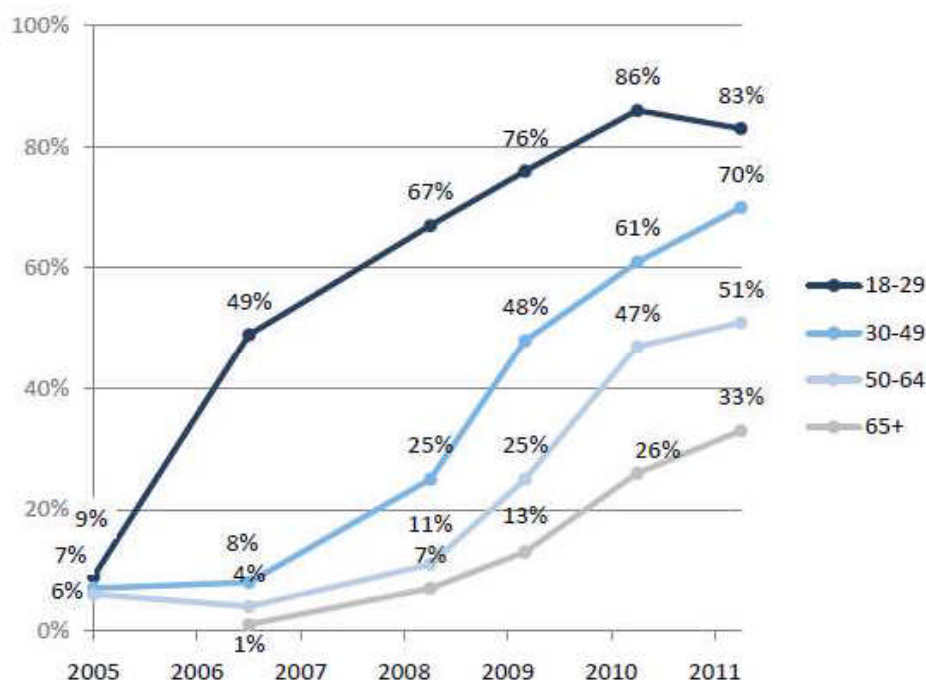
Source: Pew Research Center's Internet & American Life Project surveys: February 2005, August 2006, May 2008, April 2009, May 2010, and May 2011.

Harnessing Social Capital

Most importantly, the age of social networking users tends to be in the 18-29 age category compared with older adults, although that is changing.¹⁵

Social networking site use by age group, 2005-2011

The percentage of adult internet users in each age group who use social networking sites



Note: Total n for internet users age 65+ in 2005 was < 100, and so results for that group are not included.

Source: Pew Research Center's Internet & American Life Project surveys: February 2005, August 2006, May 2008, April 2009, May 2010, and May 2011.

As the number of people who use social networking and social media proliferates, libraries and librarians will find themselves in positions as curators of information to harness the power of social media for academic and scholarly research and reflect the demographic analysis. In Charnigo and Barnett-Ellis' 2007 survey of 126 academic librarians, they estimated 54 percent of the librarians surveyed thought there was little academic use for Facebook.¹⁶ On the other hand, in a study done approximately at the same time by the Pennsylvania State University Libraries 126 of the 441 reference questions in a sampled Fall Semester study were asked through Facebook. Compared to the other methods of reference questions such as email, phone, IM, and In-person reference, the Facebook transactions were the most highly asked for reference questions.¹⁷

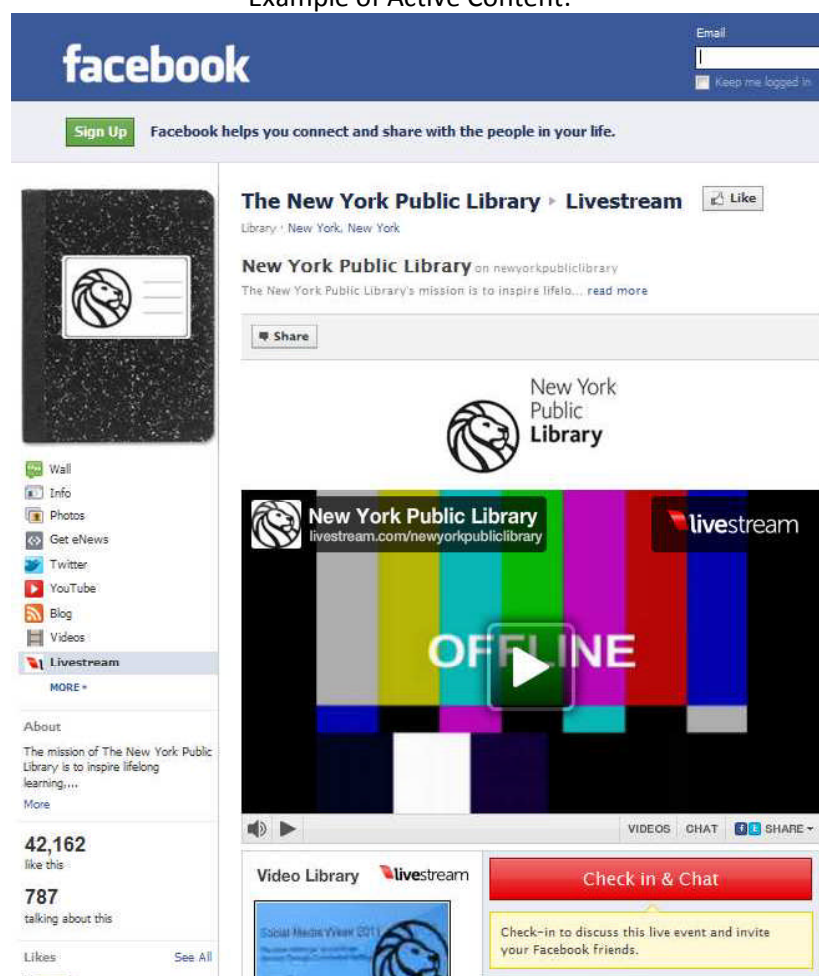
Since social media is still a relatively new concept, some libraries and librarians are resistant to the idea that social networking can have academic purpose (Dickson and Holley, 2010).¹⁸ It is assumed that the usage of this method of communication is still evolving and will continue to evolve many years to come. This paper and our research examine how libraries use social media as a product or process to achieve the goal of improving academic scholarship in different ways. In terms of academic usage, some universities, such as how George Washington University concluded in a study in 2007 that only four percent of the students that they interviewed used Facebook for academic purposes¹⁹ and we can speculate that this figure has probably increased over the last four years, but by how much remains unclear. As more campuses adopt distance education to reach out to more students, course management systems may be used to facilitate information sharing and online participation. Since Facebook has let users design programs, attach them to their own pages and publish them for other users, examples ranging from sports to entertainment, science, medicine and the arts allow users to satisfy many of their online needs without leaving their profile pages.²⁰

Library Applications – Passive/Active social media content

One of the goals of libraries branching into social media sites such as Facebook, Twitter, and Google+ is the ability for the library to conduct outreach to more users while they are not inside the library. The type of information that exists on the library social media site can be divided into two main categories: static information and active social media content. Static information is a one-way information exchange that tells the user a piece of information such as the library hours, location, and website information. Static information can exist on Facebook as a complement to the library's traditional website.²¹

Active content is a two way dialog between the library and its users. This means anything that the library does through the library social media site such as online reference services to include chat reference through Facebook or, posting a Twitter feed of a library event. Other types of active content could also include the Facebook "wall" with dialogs, postings and comments from library patrons, users, and also other Facebook members containing reviews of content, exhibitions, streaming meetings such as our gathering today among other examples.

Example of Active Content:²²



Source: New York Public Library Facebook Page – Livestream
(<http://www.facebook.com/newyorkpubliclibrary>)

Today's news that The Library of Congress is also engaged in creating and maintaining a Twitter Archive of all tweets sent suggests new social and historical value of this unique communication medium.²³

The Educause Center for Applied Research noted that the Library tops social networking in a tally of students' computer and internet activities in 2010. The list ranked 19 entries:²⁴

Activity	Share of Students Using	Median use frequency
Using College Library's website	94%	Weekly
Presentation software	93%	Monthly
Text messages	90%	Daily
Social networking websites	90%	Daily
Course or learning management systems	90%	Several times per week
Spreadsheets	86%	Monthly
Instant Messaging	71%	Several times per week
Graphics Software	67%	Monthly
Using Internet from handheld device	50%	\Daily
Voice over Internet (VOIP) from computer	47%	Monthly
Following or updating microblogs	43%	Several times per week
Contributing content to video websites	42%	Monthly
Contributing content to wikis	40%	Monthly
Video creation software	40%	Monthly
Contributing content to blogs	36%	Once per quarter/semester
Audio creation software	34%	Monthly
Online multiuser computer games	27%	Once per quarter/semester
Social bookmarking/tagging	25%	Weekly
Online virtual worlds	9%	Once per quarter/semester

Somewhat related, the same source reported that text messaging is the communication form of choice for most students compared to social networking sites, instant messaging and VOIP.²⁵ Ease of use via mobile phones, usually readily at hand suggests the probable reasons.

Organizing information appears to be central to the construct of social media and to the agenda of most libraries. One interesting sidebar may be the contributions of another computer scientist, David Gelernter, also known as a victim of the Unibomber among other notorieties. He and his students, in this context, are the chief architects of Lifestreams, defined as:

"a time-ordered stream of documents that functions as a diary of your electronic life; every document you create and every document other people send you is stored in your lifestream. The tail of your stream contains documents from the past (starting with your electronic birth certificate). Moving away from the tail and toward the present, your stream contains more recent documents --- papers in progress or new electronic mail; other documents (pictures, correspondence, bills, movies, voice mail, software) are stored in between. Moving beyond the present and into the future, the stream contains documents you *will* need: reminders, calendar items, to-do lists."²⁶

This is one of the first visible chronicles of everyday communication, exchanges and activities. Analogous to the capstone of autobiographical journal experiences, this record of sharing put document management or organization to a test for indexing, labeling or tagging.

With the subtheme of this conference, "From Social Networking to Wealth Creation," one has to consider what the different intelligence communities are. Today, the corporate sector has been described as "Data Rich, Insight Poor." This oneliner from Casper Craven, founder of the British research firm, Trovis, states, "the research warns that companies ignoring social media channels do so at their peril."²⁷

This is clearly a demonstration of the new competitive intelligence the corporate sector is relying upon -- who can utilize channeling to follow leads, connect to new customers and establish new markets. As higher education continues to emulate the business world, instruction follows business trends.

In a different domain, but one that is equally conducive to fast changes refers to the youngest population and one that the public, school and academic library community is monitoring. Children growing up today are unaware that telephones ever had cords or had to be used in a static environment. Additional studies by the Pew Foundation, Outsell, the Horizon Reports, and countless other releases from global marketing and research firms are plentiful but the report, "Zero to Eight: Children's Media Use in America" released last

month by Common Sense Media is the first examination of how “even very young children are frequent digital media users, experiencing mobile media, using computers and playing video games”²⁸ in early childhood defined as ages 2-4 and 5-8. In addition, “there continues to be a substantial digital divide, including both computers and mobile devices”²⁹ referring to family income and the ability to invest in home computers, smart phones and promise connectivity. Other findings note:

- Children under 2 spend twice as much time watching TV and videos as they do reading books
- TV continues to dominate children’s media use
- Broadcast television is the most accessible and widely used platform for educational content among lower-income children
- Media use varies significantly by race and socio-economic status, but not much by gender
- Even some young children are media multi-taskers³⁰

Evidence of media multitasking is observed as young as among 5-8 year olds with such examples as doing homework while watching television, speaking on the telephone and monitoring Facebook, but increasingly common among 12-18 year olds. Media as babysitter is also common, while out running errands, or by parents who use media to occupy themselves when they are considered to be playing with or engaging with their children. Think of the last time you visited a park or supermarket and observed parents holding a phone, wearing Bluetooth, while disciplining a child! The major conclusions confirmed:

- that increasing dosages of baby videos on TV and DVDs on a regular basis with children under 2 are concerning
- bedroom television contributes to obesity, declining school grades and other negative effects
- educational media via television remains the most popular platform
- there is an enduring and substantial digital divide among lower income families, even when targets are for health and social promotion
- quality entertainment is challenged with negative impact of content and advertising of violence or inappropriate gender roles
- perceptions of a decline in reading and literacy may not be as great as earlier studies
- there is still much to learn about neuroscience effects of multitasking by very young children³¹

Perceptions of social media characterize information products that one used to consider perpetual ways of life that were deemed safe and were part of legacies as well as future generations. As reading habits changed, information became far more abundant, accessible, yet fragmented traditions of scholarship and lifestyle have been challenged. Some academic examples include the book review. When one published a scholarly monograph or even a trade book, novel, or released a film, had an exhibition, the anticipation of the reviews clouded the authors or artists’ horizons for months as that could dictate the professional opportunities of one’s future, the commercial viability of the project, compensation packages, new merchandising potential and any host of other related by-products.

Today, reviews are immediate, distributed worldwide and can commend the work hailing its success at the cash register or box office, or pan it with a few tepid or carefully selected words. In a recent op-ed piece, a historian shared his observations of why one should read book reviews by concluding that they are “enduring value of the form,” and going on to state, “Reviews do influence my book buying –just in a roundabout way. I’m sometimes inspired to buy books by authors whose review impress me.” He concludes by saying that as a reviewer and reader of reviews, he can now see how a “pretty nifty Twitter-friendly version could have been built around the best lines from the [full review].”³² New social

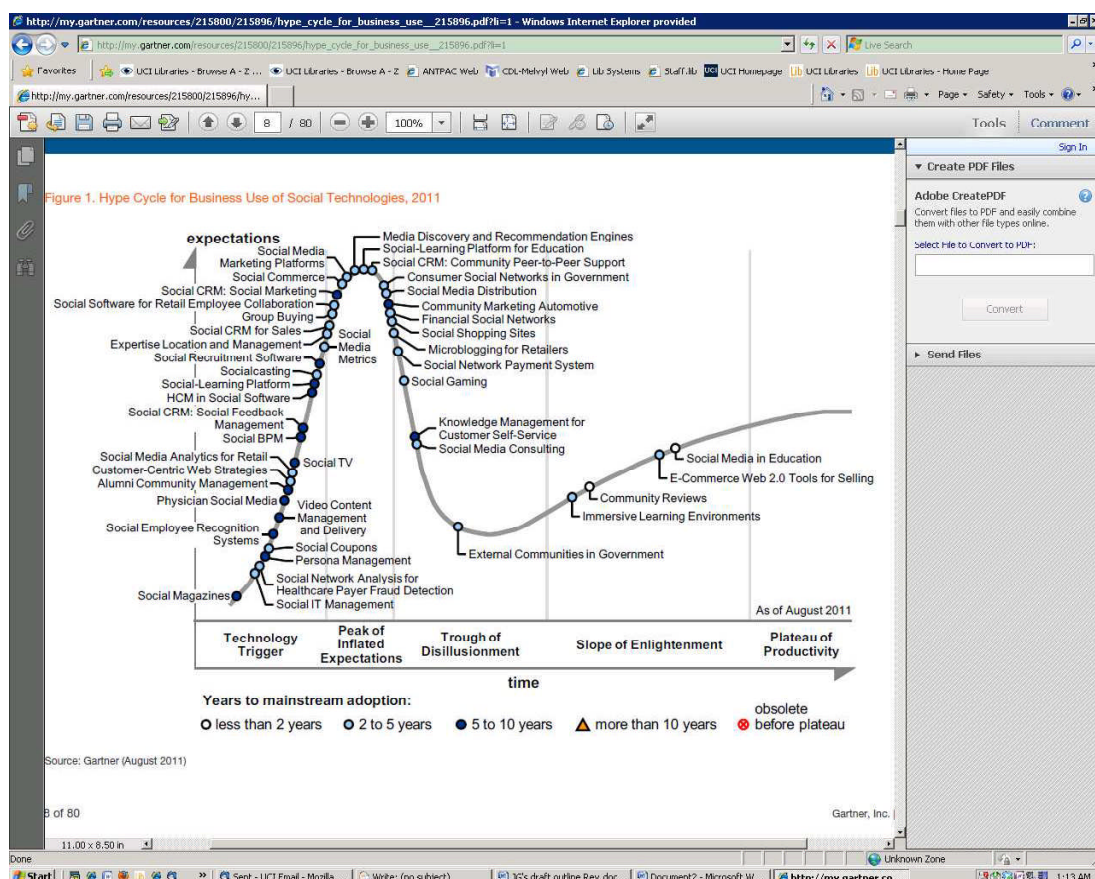
reading experiences are abundant and content ideas can be shared with people one may never encounter outside this virtual world such as in BookGlutton.com

Signs of addiction to social networking are apparent as one's identity is defined by where one is found. Earlier explorations into virtual worlds termed such visitors as geeks and techies, today, one is expected to be a joiner, self-promoter, au current with microblogging options that allow followers the 140 character tweets. Relationship building today often has its beginnings in the throes of social networking. Immediacy and constancy are what defines it as we think of the quality of the friends we acquire on these journeys.

New Communication Strategies

Forrester Researcher and Gartner are pioneers in tracking consumer behavior and indicating trends in media usage. Public opinion polling, voting patterns and behavior, campaigning, newscasts, GIS activity have each been a hallmark of grey literature. The political landscape offers a wealth of examples of grey literature that comes out of social media and could include the spectrum of that and other related matter. President Obama's 2008 campaign ignited the populace via methods of social media like nothing in the past and redefined the potential for this medium. Today's "Occupy Wall Street" again demonstrates how social media can mobilize populations with a common cause and agenda and predictably, have some staying power.

The following is an example of a Gartner Hype Cycle for Social Media in Business Technologies done in August 2011 demonstrating the adoption rate of business use of different social media:



Source: Hype Cycle for Business Use of Social Technologies, 2011. Gartner, September, 2011.³³

And yet another very academic example is the most interesting corpus of work by the political scientist, George R. Boynton at the University of Iowa who studies and teaches about new media in political discourse. He tracks the spread of political sentiments and social issues by different forms of new media, tracing how fast news travels. His research "focuses on dissemination of information of "real time" events, such as Iran's nuclear production activity and President Obama's Nobel Peace Prize. By undertaking a systematic

analysis of Twitter feeds and other social media, he sheds light on the shifting information landscape and explores emerging areas of research on the intersection between these new media and politics. Boynton publishes reports about his new media research with supporting data on his Web site, and, of course, posts updates on Twitter (@bobboynton).³⁴ For this, he was recognized as the recipient of the the 2010 Primary Source Award for Research for a project that deploys eight computers to continually harvest from the Web data on new media trends.³⁵ For an example of the Twitter Streamgraphs used to track such activity visit www.crl.edu/sites/default/files/images/focus/Boynton_stream_graph.jpg³⁶

Social Network Data

Returning to a theme of Grey Literature and data sets, the Discovery Research Group has culled some usage statistics from a variety of well-known social media platforms and shared in recent months and if anything this data has increased:

- Average Facebook user has 130 Friends
- Over 25Billion pieces of content are shared on Facebook each month
- Over 200 Million active users of Facebook via mobile phones; overall there are 500 Million users
- In late 2010, one year ago, Twitter had over 100 Million registered users with 300,000 signing up each day
- Twitter received 180 Million unique visitors to its site monthly and over 600 Million searches daily
- Twitter users are averaging 55 Million tweets daily; 637 tweets a second
- YouTube receives 2 Billion views a day and the 3rd most visited website
- On average, visitors spend 15 minutes daily viewing and adding content to site
- 24 hours of video being uploaded to YouTube each minute³⁷

As a new form of grey literature, these data sets will add to the timely corpus of information about these trends and continue to inform us about how we function in this increasingly web, semantic and multimedia age. Detecting new influences, seeking behavioral indicators, and tracking the patterns and timing of these functions and how they contribute to new innovations, stimulate the economy, generate new knowledge is all for the greater good. Higher education is testing and trying many of the features of this medium by adopting social bookmarking practices in the classroom (Brainify.com) and utilizing social media to submit college admission essays.³⁸

The Global Divergence of Social Networking depicts the growth in social media worldwide but especially in China, the US, India and Brazil, across all demographic segments. In the United States what we know is that:

- Social networking and photo sharing are the leading social channels
- Social networking and video sharing are the fastest growing platforms
- Product reviews, Q&A services and forums have the smallest differences between young and old consumers
- Despite the rise of Twitter still 50% more bloggers in the US
- Fragmentation by demographic components suggest that the target audiences and customers that use social media are not fully developed and requires special insights³⁹
- Meanwhile, the number of U.S. unique users was 50.8% December 2010, a sharp drop from 62.1% the previous June. This suggests that the use of Twitter outside the U.S. has experienced significant growth over the past six months.⁴⁰

Conclusions

Libraries, like business are trying to stay relevant and increase market shares of their constituencies. The competitive intelligence indicators suggest that engagement, branding, overcoming the digital divide and utilizing the available technologies in smart efficient ways

leads to greater participation and collaboration, key goals in the promotion of learning and exchange. Some disconnects or challenges exist for the academic community, as we are warned about the risks of creating poor content.

"If your library has been making the common mistake of primarily using social media as a marketing platform for its collections and programs, then its content probably lacks interest or relevance to the majority of its followers. One of the things that makes content successful is making sure there is a "hook" to capture readers' attention. Interesting content will generally be humorous, useful or newsworthy. Take care to ensure that whatever is being posted is not only useful but also entertaining."⁴¹

Entertaining is often perceived as inappropriate for the academic marketplace, but being relevant may require that to appeal to the target age cohort and compete with other more entertaining content. It is clear that the latest World map of social networks popularity suggests use of social media by all ages.⁴² We think it is possible, and that social networking will continue to be both product and process and dim in its hues of grey. Social Networks we speculate will continue to evolve both in the personal and external professional settings. Communication security will experience challenges but the typical features of having groups, fan pages, widgets, applications, games, profiles, communication abilities with degrees of privacy to allow for collaboration, recruitment, branding, engagement, retention, mentoring, leadership development, relationship building, promoting of diversity. With more than 90% of people in the workplace using eMail compared with fewer than 40% of workers who use social media tools, as found by a recent study commissioned by Skype, we predict a steady decline in eMail and a surge in use among a growing range of social media tools.⁴³ Users will become increasingly savvy and the unforgiving economy will lead by changes in business communication and thus education will follow. That is our scenario and grey literature will not be referred to that way because its ubiquity and increasing ease of use will continue to transform how life is conducted. When asked back in 2007 in *Friends: the Social Networking Sites for Engaged Library Services* whether social networking applications are a long term phenomenon, Stuart Weibel, Research Director at OCLC, emphatically stated, "they already ARE."⁴⁴

Let's end by asking all those in this room who over the last week (due to travel) have contributed to, consulted or read a blog, to please rise; who collaborated on or contributed to a wiki to rise; who watched a podcast to rise; who updated personal information in fb to rise; who sent a tweet to rise; who sent an eMail message to please rise. Need we say more?

References

- ¹ Ford, Rob and Wiedemann, Julius, eds., (2010). *The Internet Case Study Book*. Cologne: Taschen, 232.
- ² Zero to Eight: Children's Media Use in America (2011). *Common Sense Media Research Study*, released October 25, 2011. <http://www.commonsensemedia.org/research/zero-eight-childrens-media-use-america>
- ³ Bradley, Anthony J., (December 10, 2009). The six core principles of social media: Based collaborations. Gartner ID G00172930. Viewed October 29, 2011.
- ⁴ Global Web Index (2011). *Annual Report*. <http://globalwebindex.net/explore-the-data/data/annual-report-2011/>
- ⁵ *Ibid*.
- ⁶ Threads on LITA-L listserv, December 1, 2011. Keeping up with technology.
- ⁷ Barrett, Edward, ed. (1992). *Sociomedia: Multimedia, Hypermedia, and the Social Construction of Knowledge*. Cambridge, MA: MIT Press, 1.
- ⁸ *Ibid*, 8.
- ⁹ Rice, J. (2009). *Church of facebook: How the hybperconnected are redefining community*. Colorado Springs, CO: David C. Cook Publishers.
- ¹⁰ Lankes, R. David (2011). *Atlas of Librarianship*. Cambridge, MA: MIT Press, 31.
- ¹¹ *Ibid*, 15
- ¹² *Ibid*, 30
- ¹³ Madden, M. and Zickuhr, K. (2011, August 26). 65% of online adults use social networking sites. In *Pew Internet & American Life Project*. Retrieved October 25, 2011. <http://pewinternet.org/Reports/2011/Social-NetworkingSites.aspx>
- ¹⁴ *Ibid*.
- ¹⁵ *Ibid*.

- ¹⁶ Charnigo, L. and Barnett-Ellis, P. (2007). Checking out Facebook.com: The impact of a digital trend on academic libraries. *Information Technology and Libraries*, 26 (1), 23-34.
- ¹⁷ Mack, D., Behler, A., Roberts, B., and Rimland, E. (2007). Reaching students with Facebook: Data and best practices. *E-JASL: The Electronic Journal of Academic and Special Librarianship*, 8(2). Retrieved November 7, 2009.
- ¹⁸ Cickson, A. and Holley, R.P. (2010). Social networking in academic libraries: The possibilities and the concerns. *New Library World*, 11 (11/12), 468-479. doi:10.1108/0307480101094840
- ¹⁹ Bietila, D., Bloechi, C. and Edwards, E. (2009). *Proceedings from the ACRL Fourteenth National Conference*, 2009, Seattle.
- ²⁰ Baumann, M. (May 2008). Facebook Apps: Deeper Social Networking. *Information Today*, 25 (5), 1.
- ²¹ Dickson and Holley.
- ²² New York Public Library Facebook Page via Livestream. <http://www.facebook.com/newyorkpubliclibrary>
- ²³ Library of Congress to receive entire Twitter Archive. <http://www.federalnewsradio.com/?nid=239&sid=2658996> (December, 7, 2011)
- ²⁴ Educause Center for Applied Research (2010). *Chronicle of Higher Education*, August 26, 2011, 51.
- ²⁵ *Ibid.*
- ²⁶ Lifestreams: Organizing Your Life. <http://cs-www.cs.yale.edu/homes/freeman/lifestreams.html>
- ²⁷ Williams, Peter (October 29, 2011). Data rich, insight poor. *Information World Review*. <http://www.lwr.co.uk/social-media/3011134/Data-rich-insight-poor>
- ²⁸ *From Zero to Eight: Children's Media Use in America*, 9.
<http://www.commonssensemedia.org/sites/default/files/research/zerotoeightfinal2011.pdf>
- ²⁹ *Ibid*, 10.
- ³⁰ *Ibid*, 11-13.
- ³¹ *Ibid*, 29-30.
- ³² Wasserstrom, J. (September 9, 2011). Why read book reviews? *The Chronicle Review*, B20.
- ³³ Gartner (August 2011). Hype Cycle for Business Use of Social Technologies, 2011: 8. Gartner, Inc.: G00215896. Viewed October 30, 2011.
- ³⁴ G.R. Boynton Website. <http://www.boyntons.us/website/grb/grb-start-page.html> Viewed October 28, 2011.
- ³⁵ Center for Research Libraries (CRL) (2010). Primary Source Awards, 2010: Robert G. Boynton. <http://www.crl.edu/focus/article/6642> Viewed October 30, 2011.
- ³⁶ Twitter StreamGraph. http://www.crl.edu/sites/default/files/images/focus/Boynton_stream_graph.jpg Viewed October 30, 2011.
- ³⁷ Discovery Research Group (2011). Social Media Research and Listening. <http://www.discoveryresearchgroup.com/social-media-research/> Viewed October 30, 2011.
- ³⁸ The New College-Admission Essay: Short as a Tweet. *Chicago Tribune*, December 20, 2011.
- ³⁹ Digital Stats, April 2010. <http://digital-stats.blogspot.com> Viewed October 30, 2011.
- ⁴⁰ Sysomos (2011). Exploring the Use of Twitter Around the World. <http://www.sysomos.com/insidetwitter/geography/> Viewed October 30, 2011.
- ⁴¹ Solomon, L. (2011). *Doing Social Media So It Matters: A Librarian's Guide*. Chicago, IL: ALA Editions, Special Reports, 53.
- ⁴³ Sambhi, H., (December 2009). World map of social networks, in *Social HR*. Toronto, ON: Carswell: 119.
- ⁴⁴ Sturdevant, C. (January 3, 2011). Socializing the enterprise. *eWeek*, 28 (1): 34.
- ⁴⁵ <http://onlinesocialnetworks.blogspot.com/2007/12/blog-post.html>

Additional Selective References

- Baird, T. (2009). *How to drive traffic to your website with Facebook: Growing your online business with Facebook, and much more*. La Vergne, TN: US Emereo.
- Blossom, J. (2009). *Content nation: Surviving and thriving as social media changes our work, our lives and our future*. New York: Wiley.
- Cassidy, E.D., Britsch, J., Griffin, G., Manolovitz, T., Shen, L, Turney, L. (2011). Higher education and Emerging Technologies: Student usage, preferences and lessons for library services. *Reference & User Services Quarterly* 50 (4): 380-391.
- Gillette, F., June 27-July 3, 2011. The rise and inglorious fall of Myspace. *Bloomberg Businessweek*: 52-59.
- Katz, J.E., LaBar, W. and Lynch, E. (2010). *Creativity and Technology: Social Media, Mobiles and Museums*. Edinburgh: MuseumsEtc.
- Kear, K. (2011). *Online and social networking communities: A best practice guide for educators*. London: Routledge.
- Micek, D. and Whitlock, W. (2008). *Twitter Revolution: How social media and mobile marketing is changing the way we do business and market online*. Las Vegas, NV: Xeno Press.
- ShelfLife: Significant resources on social media. *Choice Reviews Online*, November, 2011.
<http://www.cro2.org/default.aspx?page=reviewdisplay&pids=3688423>
- Solis, B. (2011). *Engaged!: The complete guide for brands and businesses to build, cultivate and measure success in the new web*. Rev. and updated. New York: Wiley.
- Sterne, J. (2010). *Social media metrics: How to measure and optimize your marketing investment*. New York: Wiley.
- IEEE Computer Society, November 2010. Technology-mediated social participation. *Computer*, 43(11).
- Weber, L. (2009). *Marketing to the social web: How digital customer communities build your business*, 2d ed.. New York: Wiley.

Knowledge Communities In Grey*

Claudia Marzi (Italy)

Abstract

The dynamic nature of modern human social interactions, and the increasing capability of wireless and mobile devices for creating and sharing contents, open up the opportunity for a wide dissemination of information through complex knowledge sharing systems.

As the shared knowledge components build cognitive ties, there is no real sharing of knowledge without a common understanding of it.

In this article, particular emphasis is laid on technologies in Natural Language understanding and knowledge management for providing structured, intelligent access to the continuously evolving content, generated on-line in a pervasive collaborative environment.

In detail, robust automated techniques for term extraction and knowledge acquisition are used to tap the information density and the global coherence of text excerpts sampled from both general-purpose and subject-specific social networks. We show empirically that the two sources may exhibit considerable differences in terms of content accessibility and informativeness.

1. Introduction

The development of digital technologies and the continuous evolution of telecommunication networks are rapidly heading our society towards a culture of participation and to a more and more interactive communication. The dynamic nature of modern human social interactions, the adaptive networking protocols and data management systems are fostering pervasive information and communication environments.

The web represents an unlimited universe of information and data, and offers the steadily increasing availability of ubiquitous accessible information.

As accessibility improves, however, the huge amount of data and information available on the web need to be identified, classified, analyzed, filtered, so as to enhance the generation and assimilation of new knowledge.

Large volumes of information, even structured information, have to be managed, and generation and assimilation of knowledge have to be facilitated. Knowledge needs to be represented, standardized and distilled from multiple sources.

In this context, Social Networks can enhance fore-front ideas and highly innovative contents; they offer the enormous potential to transform research, and research results, into a knowledge co-creation process.

2. Methodology

2.1 Research questions

Given this scenario, the following questions arise naturally.

To what extent can Social Networks provide a real opportunity for sharing and disseminating novel information and generating knowledge?

Can they really be supportive of a steady flow of technical and scholar writing, or do they only provide a general communication channel for ephemeral communication exchanges?

Is there a specific added value in the way Social Networking can foster people's interest in sharing and building information?

Is interactive, informal and ubiquitous information exchange developing a new social framework for the creation of public-domain knowledge?

* First published in the GL13 Conference Proceedings, February 2012.

We suggest that all these questions can be addressed by applying advanced Natural Language Processing tools for automated content extraction to the analysis of web-based text collections, sampled from both general-purpose and specialized examples of social networks.

2.2 Research rationale

The Information Extraction literature provides different modes and tools for knowledge acquisition and representation: from highly structured, standardized and objective knowledge information systems based on ontological hierarchies and relations to more dynamic, subjective tools for volatile knowledge representation such as word clouds and concept maps. Technologies in Natural Language understanding offer an objective measure of the information density of a text document or document collection and ways to map out the distribution/development of information. This makes it possible to compare the information structure across texts and get a sense of their level of content sharing and knowledge coherence.

This approach will highlight current automated tools for concept acquisition and ontology learning that are conducive to an incremental access and management of content, to establish a fruitful bridge between modes of knowledge sharing/creation and dynamic, incremental approaches to automated knowledge acquisition and representation.

2.3 Methodological approach

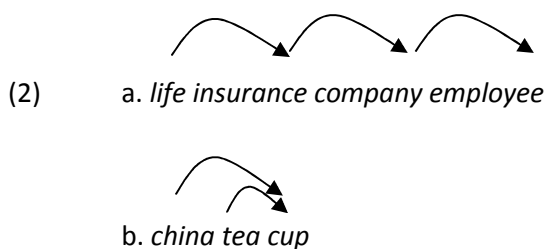
Natural Language Processing (NLP) tools can augment text documents with layers of mark-up data, making the hidden linguistic structure of the document overtly represented and accessible. The input text is segmented down into words and multi-word structures, mutually linked through syntactic relations. Moreover, salient terms are identified in context, to provide access keys to the basic contents of the document. In classical NLP architectures, this is carried out in a step-wise fashion, with layers of annotation being cascaded in a feeding relation, from *tokenized* texts to trees of dependency relations. Typical parsing steps are: i) *tokenization*, ii) *morphological parsing* and iii) *dependency trees*.

Tokenization amounts to assigning a string of characters the status of single token, where a token is the most basic parsing unit, approximately corresponding to a linguistic word, but also including non-lexical units such as dates, addresses, proper names, acronyms, measuring expressions, etc. For tokens to be identified as independent words and assigned their corresponding part-of-speech tag (or grammatical category), their set of morpho-syntactic features (e.g. number, gender, tense, etc.) and their lemma (or lexical exponent), they have to undergo a level of *morphological parsing*. In its simplest instantiation, morphological parsing requires the existence of large repositories of word forms, where each form is glossed with a set of morpho-lexical features. However, in languages with rich morphologies, a closed-list approach to morphological parsing is subject to serious risks of failure, as shown by the German example in (1) (borrowed from Anderson and Lightfoot, 2002):

- (1) *Lebensversicherungsgesellschaftsangestellter*
(life insurance company employee)

In fact, no German lexical repository can be expected to be large enough to contain all possible compounds of this kind. A principled solution is to split the compound into its simpler constituent words (Leben + Versicherung + Gesellschaft + Angestellter), for the latter to be looked up in a lexical database as individual entries.

Once word tokens are identified and categorised for contextually-appropriate part-of-speech and lexical exponence, they are grouped into larger constituents defining their *syntactic dependency relations*. A dependency relation is a binary relation linking two tokens in context, usually represented as a pointed arc going from the *dependant token* (a complement or a modifier) to its syntactic head (usually a complemented or modified verb or noun). Dependency relations can also be defined between the constituents of complex NN compounds as shown by the following examples:



The pointed arcs above tell us that *life insurance* is a type of *insurance*, *life-insurance company* is a type of *company* and *life-insurance company employee* is in fact an *employee*. Incidentally, it should be appreciated that not all dependency chains in NN compounds must look like those in *life insurance company employee*, as shown by the dependency structure of *china tea cup* above, where both *china* and *tea* entertain a modifying relationship with *cup*.

Linguistically annotated documents provide a jumping-off point for the acquisition of a more and more abstract representation of the document content, in line with the so-called “layer cake” approach to ontology learning (Buitelaar, Cimiano and Magnini 2005), whereby:

- ✓ words are structured into terms (e.g. *life insurance company* is a complex term),
- ✓ terms are grouped into conceptual classes (e.g. *insurance company* is a co-hyponym of *telecommunication company*),
- ✓ concepts are linked together through vertical (taxonomical, e.g. *life insurance company* is a hyponym of *company*) and horizontal (ontological, e.g. people are typically employed in a *company*) relations.

Such a wealth of information provides the basis to a computational platform for automated document content sharing, access and dissemination, that allows document contents to be queried by concepts and concept relations rather than by fixed text patterns or key-words (Lenci et al., 2008). For example we can search a text for information about the number of employees of a given insurance company or its overall yearly income and the like. With no linguistic information such as in (2.a-b) above, a text can be navigated only through fixed word patterns. Linguistic annotation offers a more abstract level of information which can selectively be searched for intelligent information access.

Another, orthogonal level of linguistic information that can usefully be represented through NLP technologies is the *content accessibility* of a document, defined as the level of readability of a text document calculated on the basis of its processing difficulty. Processing difficulty is a multi-factorial concept, which can be decomposed into several, fairly independent factors, such as lexical richness, lexical density and (morpho-)syntactic complexity. Each factor can be assessed independently through measurable parameters. This type of analysis allows us to get an objective measure of the information density of a text document or document collection and to map out its distribution/development through the document(s). This makes it possible to compare the information structure of different text collections and get a precise sense of their level of informativeness, content sharing and knowledge coherence.

In particular, *lexical richness* has to do with the lexicon of a text document, defined as the set of word types attested in the document. Trivially, a richer lexicon has a higher set cardinality than a poorer one. More subtly, lexical richness also involves word frequency distributions. Rare words are in fact taken to be more difficult to process than more common words and often denote the most salient pieces of content information of a document together with its level of subject-specificity. Accordingly, by inspecting the tails in the Zipfian distribution of different document lexicons we can get a flavour of the different degrees of lexical richness of the corresponding documents (or document collections). *Lexical density*, on the other hand, gives a measure of the rate at which the content of a collection is updated through the introduction of novel concepts and is defined as the number of new words that a text excerpt introduces in a document collection, divided by the length of the excerpt.

Salient domain-specific concepts and relations are most often conveyed in text through statistically significant terms. Relevant terminological units can be tracked down automatically by projecting abstract morpho-syntactic patterns such as “NP PP” (i.e. “find a syntactic structure made up out of a Noun Phrase immediately followed by a Prepositional Phrase) onto linguistically annotated texts. Text strings fitting into the targeted morpho-syntactic pattern are then filtered out through a further step of statistical post-processing, to assess their potential for termhood. Since Smadja’s (1993) seminal work, statistical methods offer reliable means of acquiring domain specific expectations concerning the joint distribution of words in sufficiently large training corpora. Association measures such as Pointwise Mutual Information (Church and Hanks, 1989) have become standard utilities to measure the degree of collocational association of word pairs in context, by exploiting the intuition that words belonging to the same bracketed pair will co-occur in corpora significantly more often than what would be expected under a model of chance co-occurrence (based on the frequency of the individual words). Based on this intuition, we can further identify the most salient terms attested in a document collection and their degree of subject-specificity, by comparing their frequency distribution in the target collection with the distribution of the same terms in a balanced, general-purpose corpus.

Finally, a score of (morpho-)syntactic complexity can be calculated on the basis of i) the average length of text clauses (the longer a clause, the more difficult to parse), ii) the way words are arranged in context (an unusual/marked order of words is more difficult than a standard word order), iii) the per-sentence length of the attested dependency chains (shorter chains are easier to be parsed and understood), and iv) the per-word distance between a head and its dependant (the longer the distance the more difficult the relationship) (Dell’Orletta et al., 2011).

This multi-factorial information allowed us to compare the information density of two different text collections: samples of text excerpts from general-purpose social networks, based on friendship relations and on social proximity, and samples of texts produced within the frame of specialized subject-based communities, based on content sharing and supporting relationships.

3. Experimental evidence

3.1 Results

Two distinct experiments were carried out on English and Italian texts. For all experiments, we used the web-based battery of NLP and knowledge-management tools made available on-line by the *Dylan Lab*¹ (Dynamics of Language Laboratory, Institute for Computational Linguistics - Italian Research Council).

In the English experiment, we conducted a cross-evaluation assessment of both grammatical and content-word parameters in three different text collections: i) a sample of messages posted in general-purpose social networks (e.g. Facebook), ii) a sample of message exchanges within subject-based web communities (e.g. LinkedIn), iii) as a base-line, a sample of Grey Literature writings (e.g. *GL 12 Conference Proceedings*). The distribution of terms in the three samples was comparatively evaluated on the basis of the degree of domain-specificity of terms. This is shown in table 1, which gives the average number of terms that were automatically found to be domain-specific by comparison with a general purpose corpus of English language newspapers (*The Wall Street Journal* section of the Penn Treebank). Characteristically, texts exchanged by subject-centred communities contain an average number of domain-specific terms comparable to the number of terms conveyed on average by GL papers, and considerably higher than the average number of terms occurring in general social network comments.

Table 1

	<i>domain specific terms (single and multiple)</i>
Social networks ²	1.75
Subject-based communities ³	14.75
GL Papers	15.75

On the other hand, table 2 gives a measure of the average syntactic complexity of our samples, resulting from the weighted integration of several quantitative parameters: lexical rarity of content-words, distribution of part-of-speech tags (an objective measure of morpho-syntactic complexity), average length of chains of dependency links stemming from a single syntactic head, and average head-complement distance measured by the number of intervening words (Dell'Orletta et al. 2011):

Table 2

	<i>difficulty level</i>
Social networks	35.30
Subject-based communities	43.85
GL Papers	55.20

Once more, the three text samples, ranked by increasing values of syntactic difficulty, reflect a gradient of content accessibility which appears to mirror the degree of communicative formality (from less formal to more formal) scored in our text types (Figure 1).

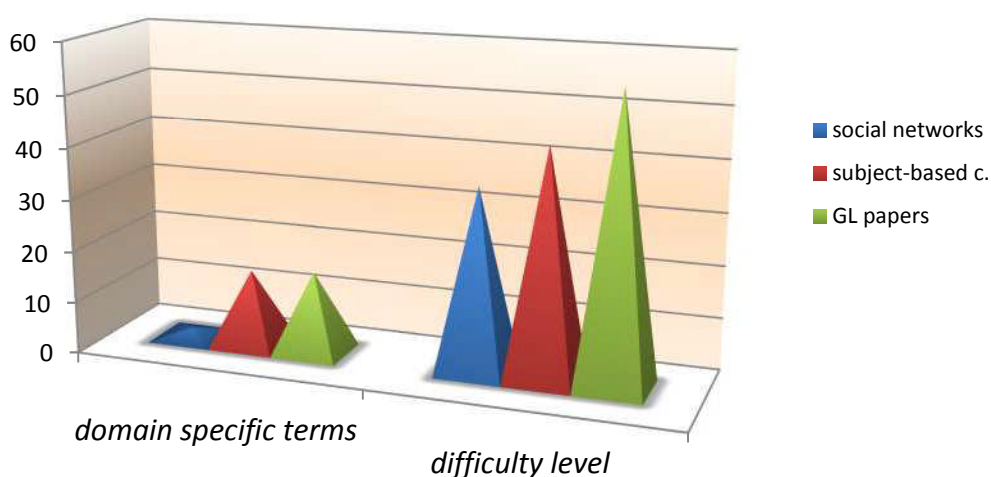


Figure 1

In the Italian experiment, we compared the overall levels of lexical coherence in two samples of Italian post exchanges: i) through a general-purpose social network (various Facebook accounts' contacts all based on friendship relations); ii) through subject-based technical blogs of a research institution (CNR intranet). Lexical coherence is automatically estimated by

measuring the flow of new lexical items that are incrementally added in a post exchange referring to the same issue. This is calculated as the number of novel words introduced by each newly posted comment divided by the length of the comment. Results are summarized in the graph below (Figure 2), providing the overall trend in the average word frequency and lexical density for the two text samples (over an exchange of maximum 20 posted comments).

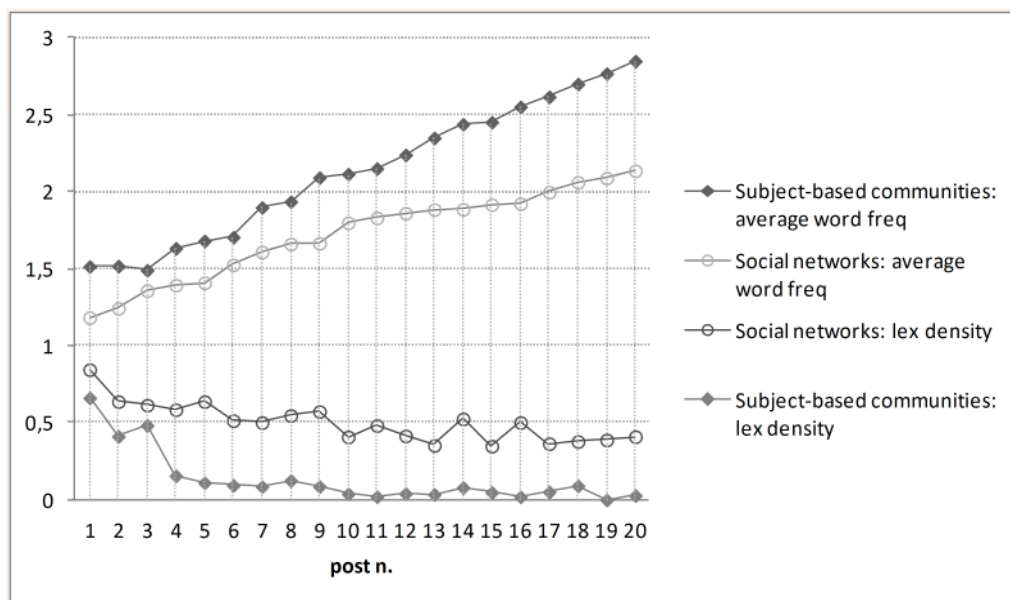


Figure 2

The two trends are remarkably different. Social networks show a slower rate of average word frequency, witnessing a higher variety of lexical choices (the same words are repeated less often). On the other hand, subject based technical writing tends to be lexically more coherent, with a systematic trend towards repetition of the same words. That these words are mostly technical is shown by the average number of domain-specific terms and their high level of difficulty/rarity as shown in table 3:

Table 3

	<i>Single terms</i>	<i>Multiple terms</i>	<i>Average domain specific terms</i>	<i>lexicon difficult level</i>	<i>Readability level</i>
Subject based CNR intra-blog	20	5	12.5	50.2	84.7
Social networks	1	3	2	12.5	73.7

3.2 Discussion

The two experiments were intended to test the empirical hypothesis that only subject-based collections offer a coherent flow of shared and structured knowledge, general purpose social networks being more erratic and ephemeral in the choice of discussion topics and domains. This hypothesis is basically confirmed by our results. The medium of Social Networks tends to make communication simpler, with shorter sentences than in traditional texts, medium/highly frequency distributed words, simpler syntax (one verb per sentence), and a high readability score. This is true of all texts that were sampled from writings exchanged through social networks, irrespective of their topic.

However, a simpler communication does not necessarily guarantee coherence of information flow and steady knowledge sharing. Only in those cases where there is a strong interaction between medium and content (as in subject-based community exchanges), we can also observe domain-specific terms, high lexical coherence, more levels of syntactic embedding, more complex readability levels. These are in fact, in our view, the hallmarks of knowledge building and informative flow.

4. Concluding remarks

NLP tools for content analysis and Information Extraction are instrumental in establishing a direct relation between modes of knowledge creation/sharing and dynamic, incremental approaches to automated knowledge acquisition and representation. They allow us to assess the content of a text in terms of its level of readability, domain-specificity, lexical coherence and density of its conceptual maps. They can be used to measure not only the effectiveness of a text in conveying information but also the extent to which this information is structured in terms of shared knowledge.

As cognitive proximity consists of sharing capabilities and knowledge in a broad context, subject based communities, primarily focused on supporting relationships and content sharing, act at the same time as providers and users of all kind of Grey Literature materials in a highly distributed and collaborative scenario, and represent a conducive environment for knowledge creation and transfer. They can represent, as collaborative networks, a key element in the advancement and dissemination of knowledge in scientific domains as well as in diverse aspects of everyday human life.

Conversely, general-purpose social networks, reflecting either friendship or superficial relationships, are virtual meeting place, but tend to generate ephemeral information and to create superficial and mosaic knowledge.

As a general concluding consideration, Social Networking is a medium with a strong potential, a house of cards powerful and delicate at the same time.

References

- ANDERSON S. R., LIGHTFOOT D. W. (2002). *The language organ: linguistics as cognitive physiology*. Cambridge University Press.
- BONIN F., DELL'ORLETTA F., VENTURI G., MONTEMAGNI S. (2010). A Contrastive Approach to Multi-word Term Extraction from Domain Corpora. *Proceedings of the 7th International Conference on Language Resources and Evaluation*, Malta, 19-21 May, 3222-3229.
- BAEZA-YATES R., RIBEIRO-NETO B. (1999). *Modern Information Retrieval*. Addison Wesley, ACM Press, New York.
- BUITELAAR P., CIMIANO P., MAGNINI B. (2005). *Ontology learning from text*. IOS Press, Amsterdam.
- CHURCH, K., HANKS P. (1989). Word Association Norms, Mutual Information and Lexicography. *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, Vancouver, Canada.
- DELL'ORLETTA F., MONTEMAGNI S., VENTURI G. (2011). Read-it: Assessing Readability of Italian texts with a View to Text Simplification. *Proceedings of the 2nd Workshop on Speech and Language processing for Assistive Technologies*, Edinburgh, UK, 73-83.
- DELL'ORLETTA F. (2009). Ensemble system for Part-of-Speech tagging. *Proceedings of Evalita'09*, Reggio Emilia, December 2009.
- FRANTZI K. T., ANANIADOU S., MIMA H. (2000). Automatic Recognition of Multi-Word Terms: the C-value/NC-value method. *International Journal on Digital Libraries*, 3(2), 115-130.
- LENCI A., MONTEMAGNI S., PIRRELLI V. (2006). Acquiring and Representing Meaning: Computational Perspectives. In A.. Lenci , S. Montemagni ,V. Pirrelli . (eds.) *Acquisition and Representation of Word Meaning. Theoretical and computational perspectives. Linguistica Computazionale, XXII-XXIII*, IEPI, Pisa- Roma, 19-66.
- LENCI A., MONTEMAGNI S., PIRRELLI V., VENTURI G. (2008). Ontology learning from Italian legal texts, in J. Breuker, P. Casanovas, M. C.A. Klein, E. Francesconi (eds.), *Law, Ontologies and the Semantic Web - Channelling the Legal Information Flood, Frontiers in Artificial Intelligence and Applications*, Springer, Volume 188, 75-94.
- MANNING C. D., SCHÜTZE H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press.
- MARZI C., PARDELLI G., SASSI M. (2011). A Terminology based re-definition of Grey Literature. *GL12 Conference Proceedings*, TextRelease, Amsterdam, 27-31.
- MARZI C., PARDELLI G., SASSI M. (2010). Grey Literature and Computational Linguistics: from Paper to Net. *GL11 Conference Proceedings*, TextRelease, Amsterdam, 118-121.
- NOOTEBOOM B. (2000). *Learning and innovation in organizations and economics*. Oxford University Press, Oxford.
- SMADJA, F. (1993). Retrieving Collocations from Text: Xtract. *Computational Linguistics*, MIT, Cambridge MA, USA, 19(1), 143-177.
- TENENBAUM J. M. (2006). AI meets Web 2.0 Building the Web of Tomorrow, Today. *AI Magazine*, American Association for Artificial Intelligence, 47-68.
- TEXT ANALYSIS TOOLS <http://www.ilc.cnr.it/dylanlab/> developed at *Dylan Lab* (ILC-CNR).

¹ Tools available at <http://www.ilc.cnr.it/dylanlab/index.php?page=software&hl=it> IT
<http://www.ilc.cnr.it/dylanlab/index.php?page=software&hl=en> US

² Facebook excerpts are messages posted during a time slot of 20 days on various accounts, varying in age, social class and professional level. The figure in the table is an average over excerpt samples whose standard deviation is 1.24

³ LinkedIn excerpts are messages taken from the GreyNet Group, during a time slot of 20 working days.

FIND THE PIECE THAT FITS YOUR PUZZLE



THE GREY LITERATURE REPORT FROM THE NEW YORK ACADEMY OF MEDICINE

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

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

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

For more information visit our website: www.greyliterature.org
or contact us at: greylithelp@nyam.org



**The New York
Academy of Medicine**

At the heart of urban health since 1847

Using Social Media to Create Virtual Interest Groups in Hospital Libraries*

Yongtao Lin and Kathryn M.E. Ranjit (Canada)

Abstract

Social networking has positively impacted the realm of human interactivity. Although libraries have traditionally been viewed as a place for research and information seeking, Web 2.0 technologies, such as blogs and Instant Messaging (IM), are starting to change users' experiences of current library services.

In this study we aim to examine whether the opportunity for enhanced researcher-librarian interaction that Web 2.0 tools provide creates a synergistic experience for health research interest groups. In addition, we will explore whether these tools increase efficiency in obtaining information and/or improve quality and quantity of research evidence.

This study was conducted by two hospital librarians working in tertiary hospitals. An email was distributed to the health care professionals from these two sites, inviting participation in two online research interest groups: Clinical Practice Guidelines and Patient Safety. A pre-survey of participants was intended to assess the comfort level of this group with Web 2.0 tools and to gauge their level of use both professionally and personally. We created a Virtual Interest Group (VIG) environment to incorporate the following Web 2.0 tools into the existing library website: user blogs, enabling commenting to facilitate a knowledge-sharing atmosphere; chat software to assist with easy access to acquired information; and Delicious tagging for a more systematic documentation of grey literature. A post-survey was conducted three months later in order to re-evaluate the participants' experiences with social media, in particular with the online interest group environment.

Findings from this study can be used to highlight future trends around the discoverability of grey literature with social media tools and to establish a basis for integrating Web 2.0 tools in library websites and services.

Introduction

Online social media, since its introduction, has transformed the means of human interaction and communication in the twenty-first century. Social media, "a group of internet-based applications that ... allows the creation and exchange of user generated content"⁽¹⁾, has advanced users' capacities in publishing information, communicating with others, and connecting and sharing with like-minded people. Social tools such as Facebook and Twitter have become dominant tools for creating knowledge, accessing information, collaborating and networking within and beyond a physical community.

Teaching and learning has often been a field where innovations in computing technology are adopted earlier than in other domains.^(2,3) Educators have, therefore, begun to explore how new paradigms of learning, such as Education 2.0, can be implemented using social media tools, including Wikis, blogs, online groups, virtual worlds, social tagging and bookmarking. These applications enable innovative behaviours in acquisition, access, processing, retrieval, presentation and distribution of information within the learning space.⁽⁴⁻⁷⁾

The impact of social media technologies is not pervasive in education; everything is affected, especially in the scholarly and research activities. Our society is increasingly interconnected through social media. Social networking can assist in developing online resources, providing a collaborative research environment. Knowledge creation and scholarly communication are moving away from a world in which a few producers generate content to transmit to a set of users. Instead, research dissemination now has various routes via a wide range of collaborative tools.⁽⁸⁾ The influence of social media on the rankings of articles has begun to be acknowledged. Binfield (2009) reviewed the metrics used by Public Library of Science (PLOS) to

* First published in the GL13 Conference Proceedings, February 2012.

analyze usage of articles appearing in their journals. They include not only page views and downloads, but also citation counts, comments, ratings, social bookmarks, and blog coverage.⁽⁹⁾ More research data is being made available online, thus greatly impacting the perception of research results. Several well-designed studies⁽¹⁰⁻¹⁴⁾ have also been published on scientists' use of Web 2.0 for information seeking of scientists, the impact of new technologies on health research, and dissemination of information through social networks.

Collective learning and social media

The advent of social media technologies imparts a new formative experience in collaborative group learning, which is often referred to as *Collective Learning*. Related to an instruction model known as active learning⁽³⁾, collective learning focuses on "the process of learning and ensures learners are cognitively engaged."^(15, p38) Collectively, the individuals can discuss, explore and develop educational content, and answer each other's questions. This enhances the engagement of learners with content, as well as their capability to actively learn and collaborate.⁽¹⁶⁾

While control is shifted from instructors to learners, undoubtedly, there are challenges to overcome. The changing roles require learners to have a strong motivation to work in groups to drive the self-directed learning process. When learning takes place through online social interaction, it challenges collaborators with sets of uncertainties and assumptions and the need to develop mutual trust between peers in a virtual research environment.⁽¹⁵⁾ Furthermore, the success of learners and the unstructured process are extremely difficult to measure.

The potential that social technologies, such as blogging, social bookmarking and social tagging, can offer foster immediate dialogue and stimulation. These applications feature the capability to support virtual social interactions and build social communities, all promising an easy-to-use, more open and dynamic platform.

The collective learning paradigm is extremely helpful in organizations⁽¹⁷⁾, where employees are expected to explore, experience, and teach themselves new technologies without formal training. Primary characteristics of social media, including accessibility of most social media sites; flexibility in refining and altering; currency and usability of these technologies help in supporting collective learning based environments. Even the social media categories (namely, *wikis, blogs, microblogging, media sharing, social bookmarking, social friendship networks and social news sites*)^(15, p40) align well with the collective learning model's goals of interactivity and proactivity. The presence of collective learning is evident in organizational learning groups such as communities of practice and virtual research environments (VRE) via social media platforms. In the context of a VRE, social networking sites such as Facebook provide an easy means of finding research partners and keeping them in touch with changing interests. Building a virtual community platform⁽¹⁸⁻²⁰⁾, where subject information guides, information gateways, literature alerts, literacy training and chat references are one-stop accessed, gives subject users a more convenient and more attractive environment to preserve their research results and to share their experience.

Social media and libraries

Libraries are major information providers and facilitators. As new communication technologies support collaboration among scientists, researchers, students, faculty members and all other users, it offers libraries and librarians the opportunity to be more proactive in facilitating the use of social networking for resource finding and sharing. Joe Fernandez (2009)⁽²¹⁾ presented a SWOT analysis of social media in libraries, allowing libraries to see the strengths and weakness that social media offer, recognize the opportunity of using social media to develop a more dynamic relationship between themselves and their users, and develop awareness of the threats that they might face when getting into the Web 2.0 realm and help them respond to the challenges accordingly.

There have been numerous studies highlighting the ongoing research activities on the integration of social media in library information and knowledge management services. The Web 2.0 tools are becoming mainstream and are increasingly more and more expected by users. For example, blogs are a popular tool used by librarians and information scientists to collaborate, connect and instruct, or to provide news and current awareness services to users⁽²²⁾, both as a form of publication and as tools for marketing their resources and events.⁽²³⁾ ⁽⁴⁾ Moreover, Bar-Ilan (2007)⁽²⁴⁾ proposed that library blogs are ideal for disseminating, commenting, and expressing opinions. Social bookmarking services such as Delicious allow users to tag (describe), save, manage and share web pages. Various library subject/resource guides are created using this social bookmarking service to facilitate the retrieval of information on the Internet.⁽⁶⁾ Information seeking behaviour within the virtual world of Second Life⁽⁵⁾ was found to be rich, complex interactive with multiple facets. Health libraries develop integrated personal information portals⁽²⁵⁾ to promote evidence-based information to health information consumers. Information literacy programs are a core service component in nearly all types of libraries, but a challenging one to meet. Learners have found themselves more engaged through interactive online learning modules created by Web 2.0 technologies.⁽²⁶⁾

Social networking and virtual interest groups (VIG)

Social networking, incorporating Web 2.0 technologies, has transformed the way information professionals are creating, storing, retrieving and disseminating information. Frameworks have been developed to facilitate the development of virtual communities.^(28,29) Virtual interest groups (VIGs), are online communities that use social networking sites to allow collaboration between users logically grouped with common shared interests.^(30, 31) Research has shown that the VIGs created and facilitated via social media platforms provide an integrated environment that supports the work of a community or group of collaborating researchers.⁽¹⁹⁾

Background

The study was conducted in two tertiary hospitals in Calgary, the fourth largest municipality in Canada and largest in Alberta.⁽²⁷⁾ The Health Information Network Calgary, a network of hospital libraries located at six major Calgary sites, provides information and library services to patients, the public and health care professionals working in the Calgary Zone and in Cancer Care. Collaboration across disciplines, geographical distances and political boundaries is increasingly accomplished in the hospitals where we provide information support.

Two key subject areas of interest that we, as librarians, have developed through our work in reference, instruction and knowledge management support are *clinical practice guidelines* and *patient safety*. Tailored information services are already in place contributing to the knowledge creation of the groups, whether through subject-based e-resource guides, or through alert services to inform researchers and health care practitioners the current literature on the related topics. Webinars and discussion forums are hosted to facilitate the sharing of information among researchers and practitioners.

Objective and Research Methodology

Our study aimed to examine whether the opportunity for enhanced researcher-librarian interaction that Web 2.0 tools provide creates a synergistic experience for health research interest groups, and to explore how the content is co-developed, assessed and further improved by users. We also hoped to gain a better understanding whether social media tools improve the discoverability of grey literature.

Emails were distributed to library users in two hospital libraries to recruit study participants for two VIGs: clinical practice guidelines and patient safety. A preliminary online survey was administered in order to understand the participants' experience using social media tools, how they used the technologies and possible barriers to using these tools for online research purposes. The survey was also intended to help the librarian facilitators identify training needs for the social media tools selected.

Blogger (www.blogspot.com) was adopted to create two blogs as a gateway platform for the other two tools integrated in our research group activities. Group members were invited to share resources and questions with their peers via this platform. Subject-related information literacy training and information consultation between researchers and librarians, and among researchers themselves were conducted via a browser-based chat service. Two platforms were explored: Meebo (<https://www.meebo.com/>) and Chatroll (<http://chatroll.com/>). Chat was used to connect with researchers virtually in real time and to enhance collaborative decision-making. Tags created for websites about clinical practice guidelines, patient safety and social media on Delicious (www.delicious.com) were linked from the two blogs. A number of networks from these tags were formed among researchers when these specified tags were subscribed to and the participants were followed.

After three months, a post-survey was distributed via email to gather feedback about the pilot project, assess comfort level with social media tools for research purposes, determine barriers to uptake during the pilot, and seek recommendations for the future services in VIGs.

Results

Pre-Survey

Twenty-five people indicated interest in participating in the VIG pilot and responded to the pre-survey. Of these 25, all elected to join the Clinical Practice Guidelines VIG and eight indicated interest in Patient Safety. Most participants had experience with at least one social media tool prior to participation in the VIGs (Table 1) and only two participants self-reported discomfort from lack of experience with social media.

Table 1. Responses of 25 participants who were asked, "What social media tools do you use?"

Tool	Number of participants	Percentage
Flickr	2	8
Twitter	6	24
Delicious	0	0
Facebook	18	72
MySpace	0	0
YouTube	13	52
Ning	0	0
Blogger	2	8
LinkedIn	2	8
Other	1	4

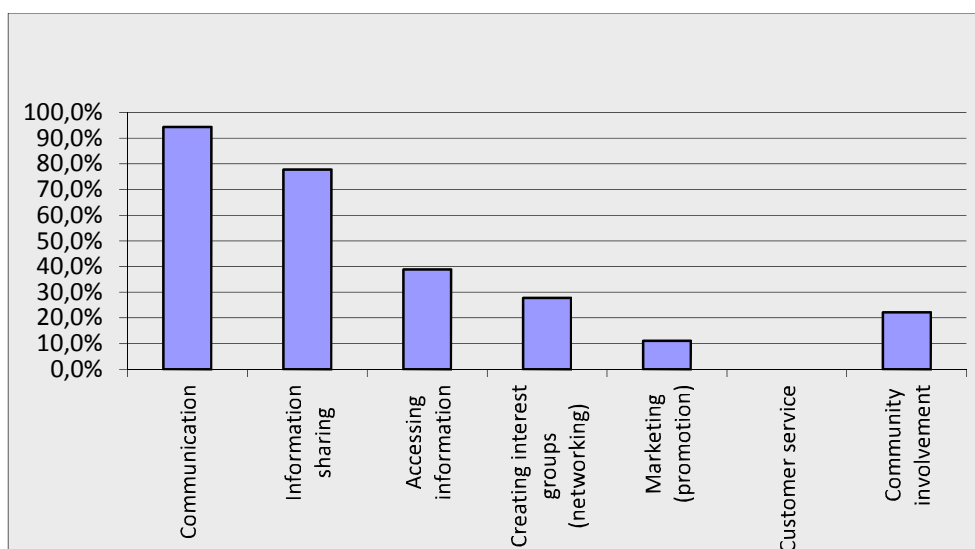


Figure 1. Participants' previous uses of social media tools.

Most participants had used social media tools for communication and information sharing purposes prior to the pilot (Figure 1).

Self-reported barriers to using social media tools included limited workplace access, time constraints, discomfort with the medium, and privacy issues. Training was requested in creating good online content and directions to use the tools proposed for the VIGs.

Blog

A blog was created for each VIG. Twenty-one participants registered for an account with one or both blogs. Twenty posts were published during the three-month period. Types of information shared in the blog posts included an introduction to each piloted social media tool and “how-to” directions, two reviews of articles pertinent to the subject areas, information about guideline review tools including AGREE-II, favourite free resources including websites and a participant’s departmental newsletter.

Chat

Four live chat sessions were held. Eight people participated in one or more chat sessions. Chat sessions were scheduled approximately every three weeks. Sessions were an open forum but each session was prefaced with questions provided to the group to foster discussions.

- Session 1: Troubleshooting issues with signing up for the blogs, why were participants interested in joining the VIGs, social media in health care and previous experience with the VIG topic
- Session 2: Social media in health care, social media use in their workplace, troubleshooting issues with Delicious
- Session 3: Reflections on recent posts, using guideline appraisal tools, exploring the use of other social media tools for research and learning purposes
- Session 4: Favourite resources in the VIG topic: websites, medical journals/databases, consultation with colleagues, etc. and why these are good resources.

Types of information shared in chat sessions included recent developments related to the interest group topic, reflections and discussions about recent posts, interest group resource sharing, discussions about the use of social media for specific information sharing purposes and new Social Media tools to explore.

Delicious

The librarian facilitators created three initial bookmark lists (tags) to share with the virtual interest groups: Clinical Practice Guidelines, Patient Safety and Social Media in Health Care. Twelve participants signed up for Delicious accounts. Participants bookmarked favourite websites during in-person training sessions. Users were encouraged to follow each others networks; however, no one was able to complete this task after the training sessions.

Types of information collected in tag lists included government websites, alert websites, online reports, statistical reports, online appraisal tools, popular reviews and videos pertaining to the interest group topics.

Post-Survey

Eight participants (32%) responded to the post-survey. Participants were asked about each tool used during the pilot, indicating how much they used each tool and assessing the value of each for research and education purposes. If the participant did not use a particular tool, they were asked to indicate why. Finally, participants were asked to characterize their whole experience in the VIG pilot.

Overall, participants reported that participation in each tool led to increased comfort level using these social media tools. Five out of eight respondents were unable to interact with one or more of the social media tools during the pilot. The greatest reported barrier to participation was time constraints (Table 2).

Table 2.

Self-reported barriers to use of individual social media tools used in the Virtual Interest Groups

	Blogs (n=2)	Chat (n=5)	Delicious (n=5)	In-person Training (n=3)
Too busy	2	5	3	3
Insufficient training			1	
Not interested				
Not useful			1	

The blog content was viewed as useful to most (five out of six) participants who read them. Chat sessions were reported as most useful for sharing and discussing interest group resources, as well as for connecting with VIG collaborators and troubleshooting issues. Of the tools introduced in this pilot, Delicious bookmarking was considered to be the most difficult tool to use. Participants reported difficulty in creating accounts for this tool and in using it to find useful resources even after in-person and online training.

Participants joined the virtual interest groups for a variety of reasons (Table 3). Most indicated the motivation to learn more about the virtual interest group topics and to connect with their colleagues over these topics.

Table 3.

Participants' motivations for joining the Virtual Interest Group pilot study

Motivation	Percentage of participants ^a
I didn't have any experience in social media so I wanted to learn some basics.	37.5%
I had some previous experience with social media and wanted to learn how social media is used for work-related purposes.	25.0%
I wanted to be involved in the discussions of the topic(s): patient safety and/or clinical practice guidelines.	62.5%
I wanted to be part of the groups to make more connections to people with interest in patient safety and clinical practice guidelines.	50.0%
I signed up because the librarian recommended the interest groups and potential benefits to me.	25.0%

^a Percentage calculated based on number of Post-Survey respondents (n=8)

Overall, participants characterized the experience as positive and reported an increased awareness of using social media tools for information sharing. Negative experiences noted were primarily due to lack of time to participate fully.

Discussion

Evolution of third interest group in social media and health care

Social Media is gaining an ever increasing role in health care.⁽³²⁾ While this pilot was designed to introduce the use of social media for learning and collaborative purposes in health care topics, it was not originally intended to address the prevalence and best practice of social media use in the health care environment. This topic, however, became a recurring theme throughout the pilot and therefore through this collective learning model the group adapted the VIGs to address questions in this topic as needed.

Social media training

The intent of this study was to facilitate as much as possible in the online environment, including training in the use of social media tools. Online learning is increasingly becoming the best practice model for continuing education in the researchers' health care environment.⁽³³⁾ However, some participants were unable to participate as desired, perhaps due to the limited time invested in online learning as each tool was introduced (Table 2). It was observed that facilitated in-person training increased the comfort level and motivation

to participate in the VIGs. The introduction of an in-person training session at the outset for learners who prefer this learning style should be considered for conducting similar projects. This introduction would likely increase participation.

Enhanced information creation and grey literature sharing in virtual interest Groups

While this pilot was small, there is strong evidence of increased information sharing and creation. Some resources that were deemed valuable by this interest group were shared in each of the social media tools used in this pilot, such as government websites, online reports and freely available research tools on the Internet. Participants from different departments and roles at three different sites came together to participate in the VIGs. The collaborative online environment allowed these professionals to share and learn from others in their organization with whom they would not usually have the chance to collaborate in-person. Information gathering took place in the use of each of the different tools. This may have been most evident in the chat sessions where questions were able to be addressed in real time. A more stable form of the shared information could be followed later in the form of a blog post or by tagging into Delicious. Examples of this can be in the sharing of departmental Twitter, which were not previously known or an organizational newsletter that was later shared in a blog post.

Future Directions

Future research will look into facilitated VIGs among health professionals and with librarians as participants. With librarians able to focus as participants and content creators instead of facilitators, the opportunity to create more awareness of grey literature in health VIGs may be more easily achievable. As more data is collected the researchers would like to explore a semantic approach of visualizing these conversations using tag-based data mapping, which will greatly improve the discoverability of grey literature in health topics.

References

- (1) Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. *Bus Horiz* 2010 2;53(1):59-68.
- (2) Johnson L, Smith R, Willis H, Levine A, Haywood K, New MC, et al. The 2011 Horizon report. New Media Consortium 2011 01/01.
- (3) Bonwell CC, Eison JA, Association for the Study of Higher Education, ERIC Clearinghouse on HE, George Washington Univ, Washington, DC School of Education and Human Development. Active learning: creating excitement in the classroom. 1991 ASHE-ERIC Higher Education Reports. ; 1991.
- (4) Goss H, Jr. Extending library instruction: using blogger to collaborate, connect, and instruct. *Journal of Library & Information Services in Distance Learning* 2010 Oct-Dec;4(4):166-184.
- (5) Ostrander M. Talking, looking, flying, searching: information seeking behaviour in Second Life. *Library Hi Tech* 2008;26(4):512-524.
- (6) Darby A, Gilmour R. Adding Delicious data to your library website. *Information Technology and Libraries* 2009 JUN;28(2):100-103.
- (7) Dlouha J, Dlouhy J. Use of Wiki tools for raising the communicative aspect of learning. 2009.
- (8) van Zyl AS. The impact of social networking 2.0 on organisations. *Electronic Library* 2009;27(6):906-918.
- (9) Binfield P. PLoS One: background, future development, and article-level metrics Proceedings of the 13th International Conference on Electronic Publishing (Milan, 10–12 June 2009) 2009:69-86.
- (10) Hughes B, Joshi I, Lemonde H, Wareham J. Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. *Int J Med Inform* 2009 10/01;78(10):645-655.
- (11) Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: Twitter and antibiotics. *Am J Infect Control* 2010 04/01;38(3):182-188.
- (12) Saver C. Tweeting, posting, and yammering: the role of social media in the OR. *OR Manager* 2010 Feb;26(2):1, 12-4.
- (13) Segbers R. Go where the customers are. Marketing (and managing) your patient experience with social media. *Mark Health Serv* 2010 Winter;30(1):22-25.
- (14) Moore A. A very sociable guy, or girl...: could social media 'branding' be a useful measure of a scientist's impact? *Bioessays* 2010 Feb;32(2):95.
- (15) Agarwal N. Collective learning: an integrated use of social media in learning environment. In: White B, King I, Tsang P, editors. *Social Media Tools and Platforms in Learning Environments*. 1st ed.: Springer; 2011. p. 37-51.

- (16) Bolhuis S. Towards active and self-directed learning. Preparing for lifelong learning, with Reference to Dutch Secondary Education. 1996 04/01.
- (17) Backstrom T. Collective learning: a way over the ridge to a new organizational attractor. *Learning Organization* 2004 01/01;11(6):466-477.
- (18) Pan W, Lan X. Building a virtual community platform for subject information services at Shanghai Jiao Tong University Library. *The Electronic Library* 2009;27(2):271-282.
- (19) Voss A, Procter R. Virtual research environments in scholarly work and communications. *Library Hi Tech* 2009;27(2):174-190.
- (20) Sonnenwald DH, Lassi M, Olson N, Ponti M, Axelsson A. Exploring new ways of working using virtual research environments in library and information science. *Library Hi Tech* 2009;27(2):191-204.
- (21) Fernandez J. A SWOT analysis for social media in libraries. *Online* 2009 SEP-OCT;33(5):35-37.
- (22) Chen C, Wu Z, Ran C, Tang Q, Chen S, Zhang X. A dynamic RSS information push service mechanism based on ontology of user information needs. *Electronic Library* 2009;27(2):222-236.
- (23) Aharony N. Librarians and information scientists in the Blogosphere: an exploratory analysis. *Library & Information Science Research* 2009 Sept;31(3):174-181.
- (24) Bar-Ilan J. The use of Weblogs (blogs) by librarians and libraries to disseminate information RID B-3452-2009. *Information Research-an International Electronic Journal* 2007 OCT;12(4):323.
- (25) Koonce TY, Giuse DA, Beauregard JM, Giuse NB. Toward a more informed patient: bridging health care information through an interactive communication portal. *Journal of the Medical Library Association* 2007 Jan;95(1):77-81.
- (26) King S, Kaplan R, MacDonald K. Teaching PubMed in cyberspace: the development of a self-learning package. *Med Ref Serv Q* 2008;27(3):272-283.
- (27) Statistics Canada. Population of census metropolitan areas 2006; Available at: <http://www40.statcan.gc.ca/l01/cst01/demo05a-eng.htm>. Accessed 1/9/2012.
- (28) Tilley CM, Bruce CS, Hallam G, Hills AP. A model for the development of virtual communities for people with long-term, severe physical disabilities. *Inf Res* 2006;11(3):55-69.
- (29) Perspectives on knowledge sharing system for special interest group (SIG): Culture, power issues and user interface. 2009 International Conference on Computer Technology and Development, ICCTD 2009; 13 November 2009 through 15 November 2009; 2009.
- (30) Kallurkar S. Targeted information dissemination 2008; Available at: <http://www.dtic.mil/docs/citations/ADA480150>. Accessed 1/9/2012.
- (31) Myhill M, Shoebridge M, Snook L. Virtual research environments - a Web 2.0 cookbook? *Library Hi Tech* 2009;27(2):228-238.
- (32) Squazzo JD. Best practices for applying social media in healthcare. *Healthc Exec* 2010;25(3):38.
- (33) Alberta Health Services. E-learning: What and how. 2010; Available at: <http://insite.albertahealthservices.ca/3645/asp>. Accessed 1/8/2012.

Information-Seeking Behaviour of Slovenian Researchers: Implications for information services *

Polona Vilar, Primož Južnič, and Tomaz Bartol (Slovenia)

Abstract

The paper presents one part of a survey of information behaviour of Slovenian researchers. Results show that Slovenian researchers in most areas show usual traits of scholarly information behaviour. Exceptions are the non-use of Web 2.0 tools for research purposes and low use of open-access materials. Survey confirmed that ICT is influential in preferences regarding resource formats, access, means for information exchange, organization of resources, writing, reading, etc. The use of grey literature is quite intensive, but depending on the academic area and research field: researchers from humanities, natural sciences and other technical fields are more inclined than social science researchers to use grey literature as the source of information for their research, and business sector is of all sectors the most concerned with patents and standards.

1. Introduction

In recent years a lot of research has been dedicated to information behaviour of various user groups, among these also to scientists. With more and more resources becoming available online, there is an evident and rapid trend towards the development of different information behaviour by scientists: what information resources they are using, how and when. Recent studies of scholarly information behaviour all show significant changes in the ways researchers communicate (Maron and Smith, 2008), publish their works (Dallmeier-Tiessen et al., 2010; Muench, 2011), collaborate (Borgman, 2009), look for information and use it (Rowlands and Fieldhouse, 2007; Information Behaviour, 2008; Palmer, Tefteau and Pirmann, 2008). The origin of all these changes is undoubtedly the impact of digital technologies.

To go with this trend we have in 2011 conducted an extensive survey of information behaviour of researchers in the Republic of Slovenia. We here present part of its results. The study as a whole aims to shed more light on the patterns of their information behaviour thus facilitating the work of research organizations, information providers such as libraries, providers of publicly funded information sources such as public research agencies. The results help to better understand the research work, evaluate current offer of information sources, as well as plan the future provision.

The segment of the results presented in this paper focusses on issues of resources and media used for research purposes. We looked at use and preferences of information sources, types of information sources used (with emphasis on grey literature), and impact of information and communication technologies on information-related activities.

2. Review of Literature

2.1. Information behaviour of researchers

Among other information related activities these studies present findings regarding scholars' use of social media, use of various information sources and of various formats. It is undoubtedly becoming clear that immense impact of the digital technologies has led to certain new patterns in scientific information and communication behaviour. Researchers have noticed that although scholars still perform some 'older' activities, such as browsing or berrypicking (Bates, 1989, 2007), they also express some previously unseen behaviours, e.g. skimming (looking at one to two pages at a time), navigating (looking around at what is available, i.e. 'the electronic sweet shop'), power browsing (reading abstracts and titles, even

* First published in the GL13 Conference Proceedings, February 2012.

indexing terms, rather than full text), squirrelling (downloading material to 'read' later), cross-checking (collecting information from different sites) (Rowlands and Fieldhouse, 2007; Maron and Smith, 2008; Information behaviour, 2008; Palmer, Tefteau and Pirmann, 2008).

In terms of resources the fact is that researchers increasingly use different resources. For some scientific communities like High Energy Physics needs immediate access to information resources and even open access to full text is not enough anymore. So the grey literature became the most important communication channel (Gentil-Becot A. 2010).

For several years in the context of e-science the initiatives of data curation, reuse of data and information, have been gaining attention. By data curation we mean management of existing research data, adding value to this data, immediate sharing of it for re-use, and long-term preservation of this data for later re-use (Choudury, 2010, Lord et al., 2004; Lord and Macdonald, 2003). This issue has been tackled by OCLC, One of the questions which arises due to the changing role of libraries serving science is not, should libraries (especially academic) undertake data curation or not, but, in what way could libraries be included in these already ongoing activities? (Gold, 2010).

2.2. Slovenian research situation and infrastructure

Development and growth of research activity in Slovenia is best shown by increase of number of research groups. In 1998 in Slovenia operated 753 research groups, in the fields of natural sciences 121 research groups, in the fields of engineering 346 research groups, in the fields of medical sciences 75 research groups, in the fields of agricultural sciences and biotechnology 60 research groups, in the fields of social sciences 98 research groups and in the fields of humanities 53 research groups. Ten years later, in 2008 in Slovenia operated 1128 research groups, which means 50% increase. In the fields of natural sciences 181 research groups, in the fields of engineering 558 research groups, in the fields of medical sciences 95 research groups, in the fields of agricultural sciences and biotechnology 80 research groups, in the fields of social sciences 142 research groups and in the fields of humanities 72 research groups.

In the year 1998 the number of registered researchers was 6971. Ten years later, this number almost doubled, with 12.182 registered researchers. The age structure of researchers was as follows: 39% of researchers were aged up to 35 years, 31% from 35 to 44 years, 19% between 45-54 years, and 11% above 55 years of age. There were 38.9% female researchers. Figure 1 shows the number of researchers, is number of FTE researchers (Figure 1). In the year 2009 on different programmes and projects (basic, applicative and postdoctoral), financed by Slovenian Research Agency (SRA) 4560 researchers have taken part (Peclin, Juznic, 2012).

Republic of Slovenia allocates around 1.5% of its BDP to support science and research. In Slovenia investment in public sector is around 100 € per capita, which puts Slovenia, relatively low, in the middle among European union countries. The results, measured by number of publications look more promising. In the year 2002 Slovenia had 809 papers per million inhabitants average of EU members was 629. In the year 2005 Slovenia had 1.104 papers per million inhabitants. EU members average was 887 and in the year 2008 the ratio 1.637 vs. 1.037. In the year 2002 Slovenia was on 9th position, in the year 2005 on 8th and in 2008 on 5th.

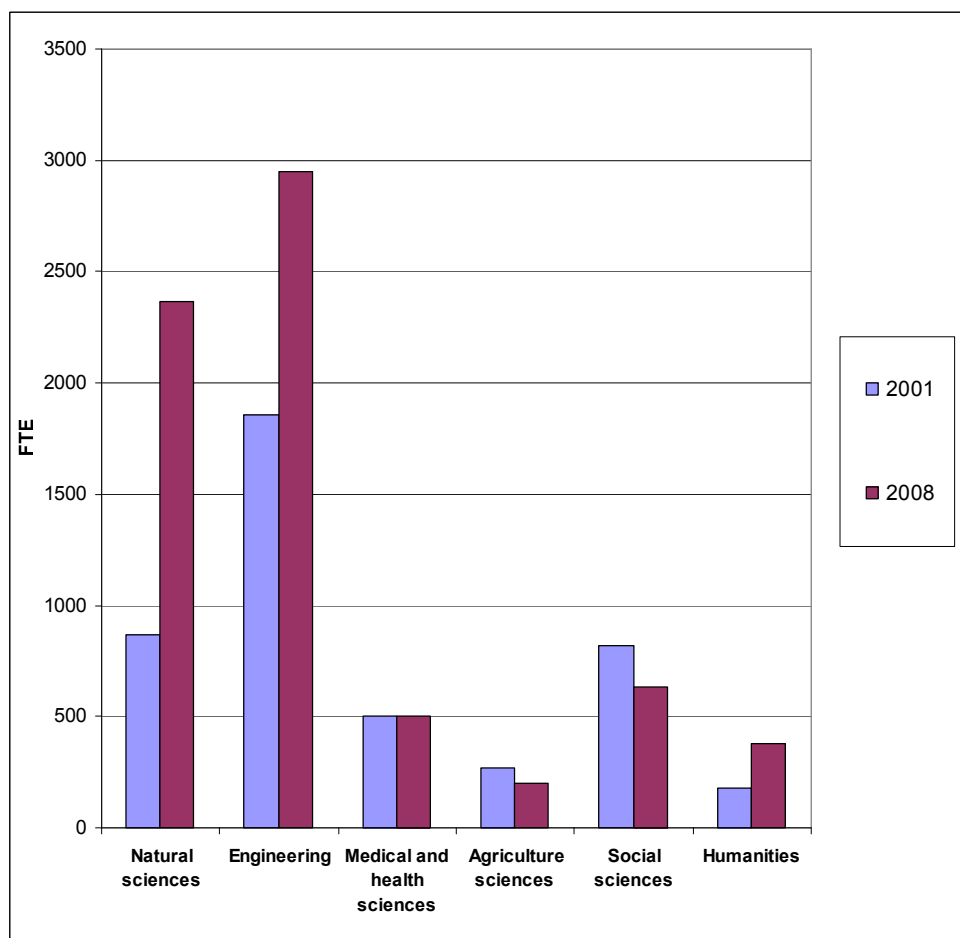


Figure 1: Number of researchers (FTE) by scientific disciplines

3. Research

3.1. Methodology

3.1.1. Sample

We used a random sample of all currently active and officially registered researchers in Slovenia. Contact details were obtained from Slovenian Research Agency (ARRS), the agency in charge of all publicly funded research in Slovenia. The agency has set up a publicly available system SICRIS (Slovenian Current Research Information System; website: <http://sicris.izum.si>) to monitor the activity of every publicly funded researcher in the country. Within this system every researcher is registered with a unique research ID number. In the year 2011 there have been around 4.800 (?) active¹ researchers. The ARRS agency provided us with researchers' contact details. Sample consisted of every eighth researcher according to allocated research ID number, which meant that there were 592 researchers who received a personal email invitation to participate in the online survey (opened from September 14th to November 14th 2011). Response rate until October 24th when we finished preparation of this text (although the poll was at that time still open) was 33.1% (196 people) of which 20.1% (119) responded adequately. Although not all questions were answered by all participants, the 119 responses still allowed a fairly detailed analysis.

3.1.1.1. Methods of gathering data

We used an online survey consisting of 25 questions: 18 content questions (Likert scale type) and 7 demographic questions. The latter seven asked participants for their gender, age, type of their current work (research, teaching), years of experience with either research or teaching, employment status (independent researcher, working at a research organization, university, commercial organization), and research area².

Content of the survey (initial 18 questions) was dedicated to various aspects of information behaviour. As mentioned, they were mostly Likert scale type questions. In this study relevant were:

- preferred format of information sources (printed, electronic, either)
- opinion on the impact of information and communication technology on various aspects of research work (searching and gathering of information sources, relevance judgement, organization of acquired sources, citation checking, reading, communication, independent writing, collaborative writing)
- number of printed and electronic sources in the personal archive
- frequency of use of information sources in the personal archive
- frequency of instances when the researcher uses electronic tools for searching for sources, but prints the sources for reading
- ratio of certain types of scientific sources which are normally cited in researcher's publications: scientific publications, electronic publications, open-source publications
- ways of acquiring scientific publications (personal subscription to printed or electronic journals, organizational subscriptions of printed or electronic journals, from e-archives or repositories, interlibrary loan or document delivery services, directly from colleagues) and the frequency of each
- attitudes and opinions concerning various aspects of handling data curation
- types of sources used in research work (eg. formal sources like books, journals, reference material, or informal (grey) sources like project reports, dissertations, social networks, blogs, forums, websites) and frequency of each
- tools used to start search for information sources for research purposes (e-journal providers, specialized bibliographic databases, specialized information portals, web search engines, library catalogue)

As explained above, every participant received an email invitation and we used the data assembled till October 24th.

3.1.1.2. Data analysis

The data was analyzed using SPSS software. Besides descriptive analysis we also performed bivariate statistics to see if any connections between demographic and content variables could be identified.

3.2. Results

3.2.1. Characteristics of respondents

There were 46,5% female researchers. The age structure of respondents was as follows: 27.6% were aged from 20-30 years, 36.7% 31-40 years, 17.3% 41-50 years, 12.2% 51-60 years and 6.1% above 60. The structure of respondents according to research area is shown in Table 2. Of the 119 respondents who are all, as explained, active researchers, 91 are involved in research, 60 in teaching and 16 in other activities (more than one current activity is possible for an individual; for example, besides being researchers some are also university teachers, medical doctors, etc).

Table 2: Structure according to research area³

	No.	%
Natural Sciences	30	25.2
Technical Sciences	20	16.8
Medicine	12	10.1
Biotechnology	8	6.7
Social Sciences	21	17.6
Humanistic Sciences	15	12.5
Interdisciplinary Research	13	10.9
All	119	100

We can say that they are rather experienced both in research and in teaching. The data can be seen in Table 3. In terms of their employment status, 3.4% work as independent researchers, majority are employed either at university (52.9%) or some type of public research institute (32.7% in research organizations which are part of some institution and 5.9% in independent research organizations), 2.5% work in economy and 4.2% in other organizations, such as hospitals (again, one individual can be employed in one or more institutions). When asked about their current work tasks (multiple answers were possible), 91.9% answered that they do research, and 60.6% that they also teach – from this we can see that two thirds are actually involved in both activities.

Table 3: Experience in research and teaching.

	Research experience		Teaching experience	
	No.	%	No.	%
Less than 1 year	1	1.0	16	19.3
1-5 years	29	29.6	27	32.5
6-10 years	24	24.5	13	15.7
11-15 years	15	15.3	9	10.8
Over 15 years	29	29.6	18	21.7

3.2.2. Information behaviour

3.2.2.1. General traits of information behaviour

Generally speaking, we found that Slovenian researchers do not differ to a great extent from what has been discovered for today's scientists. In the first section we present some general traits of their information behaviour. If we look at the use of types of information sources – with this we mean the ways researchers come to information they need for research purposes – (Table 4), we see that researchers intensively use formal sources (which is, of course, understandable). However, also the use of informal sources (grey literature) is quite strong. Over one third of researchers (35.3%) use them always or often, and another fourth (24.2%) occasionally. Research reports and dissertations are used often by one fourth (25.3) and occasionally by half (51.5%). Over one fourth (27.5%) often or always acquire information from e-archives, if we add those who use them occasionally (38.5%), this is more than two thirds.

Personal contacts are another strong source of information. From Table 4 it is evident that communication produces a lot of research-related information. It is evident that researchers communicate in all fields: with colleagues within their own organization, with Slovenian colleagues, and with colleagues abroad. Intense contacts with foreign colleagues indicates how aware are Slovenian scientists of importance of international contacts and of international character of science in general. Table 4 also shows that they very often use colleagues to acquire resources (articles). 23,6% do this often or always and another 47.3% occasionally.

Table 4: Types of information resources used and their frequency

Resource type	Frequency of use (%)				
	Never	Almost never	Occasionally	Often	Always
Printed books	2.0	11.1	36.4	30.3	20.2
E-books	3.0	12.1	48.5	24.2	12.1
Printed journals	0	21.2	34.3	28.3	16.2
E-journals	0	3.1	12.2	43.9	40.8
Reference sources	5.1	20.4	46.9	19.4	8.2
Patents, standards, reports	22.2	41.4	24.2	11.1	1.0
COBISS/OPAC (Slovenian union cat.)	0	13.1	33.3	34.3	19.2
Bibliographic databases	10.1	14.1	28.3	33.3	14.1
Raw data sources	39.8	26.5	15.3	15.3	3.1

Proceedings	4.0	23.2	40.4	27.3	5.1
Preprints	9.1	41.4	29.3	20.2	0
Reviews	14.1	44.4	30.3	11.1	0
Research reports, dissertations	2.0	21.2	51.5	25.3	0
Communication with colleagues in own org.	2.0	12.1	37.4	40.4	8.1
Communication with colleagues in Slovenia	4.0	22.2	48.5	19.2	6.1
Communication with colleagues abroad	4.1	20.4	46.9	20.4	8.2
Social networks	65.7	19.2	12.1	3.0	0
Forums, disc. Groups	28.3	36.4	23.2	11.1	1.0
Library	16.5	18.6	41.2	17.5	6.2
Email alerts	19.2	18.2	34.3	21.2	7.1
Blogs	51.5	31.3	10.1	6.1	1.0
Invisible college (conferences, meetings, etc.)	3.0	19.2	44.4	27.3	6.1
Web portals	30.6	29.6	24.5	12.2	3.1
Websites	5.1	14.1	41.4	32.3	7.1
E-archives	17.3	30.6	33.7	14.3	4.1

Table 5: Ways of acquiring resources and their frequency

How resources are acquired	Frequency (%)				
	Never	Almost never	Occasionally	Often	Always
Personal subscription of printed journal	63.3	17.4	11.9	6.4	0.9
Personal subscription of e-journal	68.8	18.3	10.1	2.8	0
Organizational subscription of printed journal	7.3	21.8	33.6	24.5	12.7
Organizational subscription of e-journal	5.5	1.8	16.5	45.0	31.2
E-archive, repository	13.8	20.2	38.5	23.9	3.7
Interlibrary loan	13.6	36.4	39.1	10.0	0.9
Colleagues	5.5	23.6	47.3	20.9	2.7

As for library services, we can say that researchers are not very enthusiastic about them. Table 5 shows that the use of library among researchers is poor (41.2% use it occasionally, 35.1% never/almost never), as is the use of ILL for acquisition of resources (Table 5 tells us that 13.6% never use it, and 36.4% almost never use it). However, we were rather pleased to notice that some library services are rather well accepted: 53.5% use OPAC often/always; 44.8% often/always start search with OPAC (Tables 4 and 6).

Table 6: Information resources used to start research-related searches

Resource	Frequency of use (%)				
	Never	Almost never	Occasionally	Often	Always
E-journal sites (eg. Science Direct, Sage,...)	3.2	8.4	27.4	33.7	27.4
Specialized bibliographic databases (eg. Medline, Inspec,...)	9.4	12.5	25.0	33.3	19.8
Information portals, cross-search engines (eg. DiKUL – search portal of UL)	28.1	31.3	24.0	11.5	5.2
Web search engines (eg. Google)	2.1	9.3	11.3	45.4	32.0
COBISS/OPAC (Slovenian union cat.)	4.2	20.8	30.2	31.3	13.5

3.2.2.2. About data curation and use

Researchers were also asked to express their opinions regarding some aspects of data curation (Table 7). They are very much in favour of having the data available (71.7% agree or strongly agree) and are also prepared to provide their own research data for curation and re-use (72.8% agree or strongly agree). But, they have doubts concerning ethical dilemmas: 58.1% agree or strongly agree that it could be questionable to provide data for re-use, since it could lead to theft of results, misuse of data, etc. This is also evident when they express opinions specifically about their own research area; although they mostly think their area would benefit from this, a significant portion also thinks it would be questionable. Interesting is their opinion about libraries playing a role in data curation. The first question shows quite strong opposition to this idea, but the second, control question proved that they would also support it.

Table 7: About data curation

Issues of data curation	Frequency (%)				
	I don't agree at all	I don't agree	I agree nor I don't agree	I agree	I agree completely
Data should be available to other researchers	3.3	9.8	15.2	40.2	31.5
I am willing to provide my research data	5.4	9.8	12.0	47.8	25.0
There are ethical dilemmas regarding data curation & use (eg. theft of results, misuse of data)	7.5	14.0	20.4	44.1	14.0
In my area data re-use is questionable	8.0	43.2	21.6	20.5	6.8
My area would benefit from data curation & re-use	0	9.7	21.5	40.9	28.0
Libraries should not be involved in data curation	4.3	10.1	18.8	31.9	34.8
Libraries should be involved in data curation	6.9	12.6	27.6	27.6	25.3

3.2.2.3. Impact of information and communication technologies

Already from the data presented earlier it is evident that ICT has had a profound influence on Slovenian researchers and on their information behaviour (which, of course, does not differentiate them from researchers elsewhere): researchers like electronic materials and use them widely. They are strong users of web search engines (77.4% use them often or always), and websites (39.4% of frequent or regular users). This on the one hand again proves that researchers in some areas behave much like general public. On the other hand, the use of e-journals clearly characterizes them as scholars (61.1% use e-journals often or always). E-preferences are also evident from other data:

- 49.6 prefer to have resources in electronic format (compared to 5% who prefer printed resources),
- 51.3% have over 200 electronic papers in their personal archive,
- 38.1% of researchers cite 81-100% e-resources in their publications.

We also asked researchers how often they print electronic materials when they want to read them: the answer was that 50.0% often and 14.0 always do that.

When asked to estimate which areas of their work were made easier or harder by ICT, results (presented in Table 8) showed that a great majority more easily searches and acquires resources (99%), organizes them (83.5%), chains citations (91.3%), writes (in collaboration (84.8%) seems to be even a little easier than alone (71.9%)) and, of course, communicates (93.9%). Areas which are more difficult for many are relevance judgement (23.7%) and reading (25%). Very few think that ICT doesn't have any influence.

Table 8: ICT makes easier/harder

Activity	Frequency (%)				
	Much easier	Easier	No change	Harder	Much harder
Search & acquisition	8.0	19.1	0	0.9	0
Relevance judgement	22.8	40.4	13.2	19.3	4.4
Organization	48.7	34.8	13.0	3.5	0
Citation chaining	54.8	36.5	6.1	2.6	0
Reading	6.9	28.4	39.7	22.4	2.6
Communicating	61.4	32.5	5.3	0.9	0
Independent writing	26.3	45.6	22.8	2.6	2.6
Collaborative writing	45.5	39.3	12.5	1.8	0.9

3.2.2.4. Some surprising findings

In some areas of information behaviour Slovenian researchers managed to surprise us. The use of Web 2.0 tools for research purposes is almost non-existing: social networks are never or almost never used by 84.8%, weblogs are obviously not considered an information source – as 82.8% don't use them, and, similarly, web forums are never or almost never used by 64.6% of respondents (Table 4). Also, it seems that quite a significant proportion of researchers (20%) don't use (and cite) electronic resources very often when they write (Table 9).

When we look at other information resources (Table 4), we see that half never or almost never use preprints and a good third similarly never or almost never uses email alerts. Cross-search services and specialized portals (Table 5) seem to be perceived as equally unimportant (60.2% never or almost never start their search using those). Interestingly, also open-access materials don't seem to be very popular, as data shows that they are not used to a great extent – open-access materials comprise less than 20% of citations for 58.3% of researchers (Table 9). All these findings differ from the usual findings for contemporary scientists and would be worth looking into.

Table 9: Shares in citations⁴

Shares	Frequency (%)				
	Below 20%	21-40%	41-60%	61-80%	81-100%
Share of scientific publications	3.7	2.8	10.1	20.2	53.8
Share of open-access publications	58.3	15.6	13.5	6.3	6.3
Share of electronic sources	20.0	14.3	7.6	20.0	38.1

3.2.2.5. A closer look: Information behaviour in relation to demographic variables

We performed bi-variate statistics (Chi-Square tests) to investigate some connections between variables. A link can be identified between age of respondents and their perceptions of various aspects of digital tools, formats, ways of communicating and acquiring information, younger researchers tend to prefer digital environment.

Gender of respondents is not an issue which would influence preference or behaviour. This is, of course, not surprising, since gender has not been identified as a factor influencing scholarly information behaviour. However, we did identify research discipline as being influential in some respects:

- Natural Sciences:

- are keen to use research papers, dissertations,
- don't acquire their resources by print journals or interlibrary loan,
- use raw data,
- cite higher proportion of scientific literature.

- Social Sciences:

- cite higher proportion of scientific literature,
- don't think libraries should undergo data curation.

- Technical Sciences:

- don't use raw data,
- use standards, patents,
- are in favour of libraries doing data curation.

- Humanistic Sciences:

- use research papers, dissertations,
- prefer printed sources,
- acquire their sources by using print journals,
- cite higher proportion of scientific literature, lower share of e-sources and lower share of open-source materials,
- are not keen on using ICT to organize resources.

- Interdisciplinary Research:
 - use e-archives to acquire their resources
 - are in favour of data curation from various aspects: that data should be available to others, that they would provide their own data and that libraries should deal with data curation
- Medicine:
 - use websites,
 - use colleagues to acquire their resources and information (invisible college),
 - use raw data, but are sceptical in terms of ethical dilemmas of its use,
 - think that ICT helps independent writing.
- Biotechnology:
 - use research papers, dissertations,
 - prefer e-sources,
 - support availability of raw data,
 - are keen on using ICT to communicate and organize resources.

Other influential demographic variables seem to be employment status, experience in teaching and researching, and current job tasks.

4. Discussion

Our results in this preliminary analysis clearly show that Slovenian researchers in most areas don't differ significantly from what has been discovered for today's scientists. There are some areas which show some discrepancies – primarily the use of Web 2.0 tools for research purposes. While it was clearly shown that researchers strongly lean on personal contacts (i.e. social networks), this has for some reason not transferred to digital environment. Digital scholarship has in recent years been gaining importance in the professional lives of scientists, both for research and for teaching (see for example Veltsianos and Kimmons (2012), *Researches of tomorrow* (2011), Moran, Seaman and Tinti-Kane (2011)). Non-use of weblogs and digital social networking tools can to a large extent be contributed⁵ to work overload (most researchers also teach), lack of motivation and maybe also as a reaction to popular use of social network tools in today Slovenia, not so much to technological issues. Last political elections that taken place and used social network tools as a political parties promotion tools. Some stated that currently they are 800.000 facebookes in Slovenia, which is a lot for two million of Slovenians, as 60% of Slovenians have their profiles (63% are females) (RIS, 2011). Slovenia is, in fact, one of the most typical European Union (EU) countries with respect to information technology usage (Petrič et al. 2011). Some current research imply that it appears that many users are mainly using Facebook to partake in passive activities, instead of providing active social contributions (Ryan, Xenos, 2011). Such findings suggest that not all Facebook users are using the site to improve their social capital, which can also some impact on nonuse.

Another issue which has obviously not gained a lot of researchers' attention is open-access. We plan to look into these issues more deeply, to be able to shed more light on our results. One of the reasons might also be the traditionally well-organised access to academic journals in Slovenia (Juznic, 2009).

Otherwise there is no doubt that information and communication technologies have a big impact on the life and work of scientists, regarding many issues: resource formats, access, means for information exchange, organization of resources, writing, reading, etc. We also saw that, regardless of the quality of infrastructure, researchers are, as always, independent, innovative and creative to find ways to acquire information which they need and to use it appropriately.

In some respects Slovenian researchers are (much like researchers in general) quite similar to general public – perhaps more than they would like (or more than providers of formal information sources would like). This is proven by e.g.:

- Intensive use of web search engines, which is much more common than the use of 'formal resources', such as e-journal providers or specialized bibliographic databases.

- Preference of e-materials and tools, immediate access to full-text,
- Dissatisfaction with existing library services and non-use of these services (with exception of OPAC).

There are, of course, areas which clearly distinguish scientists from general public. One of those is undoubtedly reference judgement. As proven also by our results, scientists are (for obvious reasons) much more than average users concerned with judging the content and quality of their information sources.

As for grey literature, we saw that it is quite intensive, but depending on academic area and sector of employment:

- Technical, natural, humanistic and biotechnological sciences are more inclined to use grey literature as the source of information for their research.
- Business sector is of all sectors the most concerned with patents and standards.

When we asked researchers what they think about data curation, the answers were positive, but careful: in principle they are in favour of it, but many have dilemmas regarding ethics. It is also quite obvious that they (with exception of medicine) don't use these resources.

5. Conclusion

Our research has shed more light on information behaviour, and to a certain extent, also information needs of Slovenian scientists. When the study, which is partly presented here, is finished, we hope that the picture will be even more complete. Some interesting issues arose which we intend to investigate further.

What implications could our findings have for information services to scientists? Libraries will need to further rethink the services which they offer to scientists. There is no doubt that formal resources will still retain their value and importance, however, they will be shaped by information and communication technologies and new patterns of scientific publishing (e.g. open access, e-archives, repositories). Some library services are already less important to scientists – hence the decline of library visits – while others, like OPACs (with access to full-text) and provision of access to e-journals, retain their value. There are other interesting issues which libraries might consider taking up: setting up e-archives and repositories, approach data curation.

Undoubtedly formal information tools need further development in terms of more intuitive interface design, change of record structure - and that is potentially another big task for libraries. For this, of course, wider co-operation is needed, e.g. with search engine developers, publishers.

We must mention that our results were obtained with web survey and are thus a reflection of self-reported activities and decisions of scholars. Due to the limitations of this method the results should be crosschecked using some other methodology. For example, observation of scholars at work would probably give us more detailed insight into their activities, and follow-up interviews would certainly shed more light into the context of their work, on their decisions and on reasons for their actions. However, it would be very difficult, if not impossible, to assemble so many cases as we did using our current methodology.

6. References

- Bates, M.J. (1989), "The design of browsing and berrypicking techniques for the online search interface", *Online Review*, Vol. 13 No. 5, pp. 407-24.
- Bates, M.J. (2007), "What is browsing—really? A model drawing from behavioural science research", *Information Research*, Vol. 12 No. 4, available at: <http://InformationR.net/ir/12-4/paper330.html> (accessed October 10, 2011).
- Borgman, C.L. (2009) "Scholarship in the Digital Age: Blurring the Boundaries between the Sciences and the Humanities (Keynote Talk)" *Digital Humanities Conference*. College Park, MD. Jun. 2009. Available at: <http://works.bepress.com/borgman/216> (accessed October 10, 2011)
- Chowdhury, S. (2010). "Data curation: an ecological perspective", *College & Research Libraries News*, vol. 71 no. 4 (April 2010), pp. 194-196.

- Dallmeier-Tiessen, S., Goerner, B., Darby, R., Hyppoelae, J., Igo-Kemenes, P., Kahn, D., Lambert, S., Lengenfelder, A., Leonard, C., Mele, S., Polydoratos, P., Ross, D., Ruiz-Perez, S., Schimmer, R., Swaisland, M., van der Stelt, W. (2010). "Open Access Publishing - Models and Attributes", The SOAP Consortium, 2010. Available at <http://edoc.mpg.de/get.epl?fid=71514&did=478647&ver=0> (accessed October 10, 2011)
- Gentil-Beccot, A. (2010). »The driving and evolving Role of Grey Literature in High-Energy Physics.« In D. J. Farace & J. Schöpfel (eds.). *Grey Literature in Library and Information Studies*. De Gruyter Saur, Berlin/New York
- Gold, A. (2010). "Data Curation and Libraries: Short-Term Developments, Long-Term Prospects", available at http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1027&context=lib_dean (accessed November 11, 2011)
- Information Behaviour (2008), "Information behaviour of the researcher of the future – Executive summary", available at: www.jisc.ac.uk/media/documents/programmes/reppres/gg_final_keynote_11012008.pdf (accessed October 10, 2011).
- Južnič, P. (2009). "Using a decision grid process to evaluate library collections and e-journals. *New libr. world*, 110 (7/8), 341-356.
- Lord, P., Macdonald, A. (2003). "E-Science Curation Report: Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision", available at http://www.jisc.ac.uk/uploaded_documents/e-ScienceReportFinal.pdf (accessed 22 October 2011)
- Lord, P., Macdonald, A., Lyon, L., Giarretta, D. (2004). "From Data Deluge to Data Curation", available at <http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/150.pdf> (accessed October 22, 2011)
- Maron, N. and Smith, K.K. (2008). "Current Models of Digital Scholarly Communication: Results of an Investigation Conducted by Ithaka Strategic Services for the Association of Research Libraries", available at: <http://www.arl.org/bm~doc/current-models-report.pdf> (accessed October 10, 2011).
- Muench, V. (2011). "Open Access shaking the basics of academic publishing", Online, Jul/Aug 2011, pp. 18-21.
- Moran, M., Seaman, J., & Tinti-Kane, H. (2011). "Teaching, learning, and sharing: How today's higher education faculty use social media for work and for play", available at <http://www.pearsonlearningsolutions.com/blog/2011/05/09/teaching-learning-and-sharing-how-todays-higher-education-faculty-use-social-media/> (accessed December 22 2011).
- Palmer, C.L., Tefreau, L.C. and Pirmann, C.M. (2009). "Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development", available at: www.oclc.org/research/publications/library/2009/2009-02.pdf (accessed October 10, 2011).
- Pečlin, S. Južnič, P. (2012). »Research in the fields of medicine in Slovenia – research potentials, funding and publications.« in press.
- Petrič, G., Petrovčič, A., Vehovar, V., (2011) Social uses of interpersonal communication technologies in a complex media environment. *European Journal of Communication*. 26(2) 116–132
- Researchers of Tomorrow: A three year (BL/JISC) study tracking the research behaviour of 'Generation Y' doctoral students: Second Annual Report 2010-2011. May 2011. Available at <http://www.efc.co.uk/projects/rot/Researchers%20of%20Tomorrow%20Year%202%20report%20final%20110510.pdf> (accessed December 22, 2011).
- Ryan, T., Xenos, S. (2011), Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. *Computers in Human Behavior* 27 (2011) 1658–1664
- Rowlands, I. and Fieldhouse, M. (2007), "Information behaviour of the researcher of the future", Work Package I: Trends in Scholarly Information Behaviour, available at: www.jisc.ac.uk/media/documents/programmes/reppres/ggworkpackagei.pdf (accessed October 10, 2011).
- Schöpfel, J., Stock, C. (2009). Grey literature in French digital repositories: a survey. In GL10 Conference Proceedings. Tenth International Conference on Grey Literature: Designing the Grey Grid for Information Society. Amsterdam. Available at: http://archivesic.ccsd.cnrs.fr/docs/00/37/92/32/PDF/GL10_Schöpfel_Stock_final.pdf (accessed October 10, 2011).
- Veletsianos, G. and Kimmons, R. (2012). "Networked participatory Scholarship: emergent techno-cultural pressures toward open and digital scholarship in online networks", *Computers & Education*, vol. 58, pp. 766-774.
- Williams, P. and Rowlands, I. (2007), "Information behaviour of the researcher of the future". Work Package II: The Literature on Young People and their Information Behaviour, available at: www.jisc.ac.uk/media/documents/programmes/reppres/ggworkpackageii.pdf (accessed October 10, 2011).

Acknowledgment and funding

This research was conducted as a part of the project V5-1016 funded by the Slovenian Research Agency.

¹ Active means that in the current year 2011 the researcher has been allocated at least 100 publicly funded research hours.

² According to the classification of Slovenian research agency (ARRS): 1 – Natural Sciences, 2 – Technical Sciences, 3 – Medicine, 4 – Biotechnical Science, 5 – Social Science, 6 – Humanistic Sciences, 7 – Interdisciplinary Research.

³ As explained in the introduction, we used the classification of the Slovenian Research Agency (ARRS).

⁴ With citations we mean resources which are cited by researchers in their own publications.

⁵ At this point we can only provide anecdotal evidence, personal experience and results of ad-hoc interviews with individual scientists from various areas – but we intend to look into these issues more deeply in the future.

Blog Posts and Tweets: The Next Frontier for Grey Literature*

Marcus Banks (*United States*)

Introduction

My interest in grey literature began as an Associate Fellow of the US National Library of Medicine (NLM) from 2002-2004. Many colleagues at NLM and throughout the country work to improve information access for people in the public health workforce. In comparison to resources available for practicing clinical medicine, the information needs for public health are more diffuse and often require access to grey literature ^[1]. *In response to these needs NLM has developed the Partners in Information Access for the Public Health Workforce portal, which includes some avenues to grey literature but mostly useful links and validated search strategies for PubMed* ^[2]. *In addition to NLM's Partners page, staff of the New York Academy of Medicine Library has maintained the Grey Literature Report for several years* ^[3]. *This is a portal to documents produced by reputable organizations in public health and health policy. During my NLM fellowship years, debate about open access publishing - specifically, how to secure access to publicly funded research - was heating up. An open access publication is generally a type of white literature, and is available for free online and stored in a digital repository* ^[4]. *Debate over the proper balance between open and subscription access will continue for years, as library associations and publishers continue to hire lobbyists and issue strongly worded statements. In 2004 I was certain that pure open access would prevail, but now think a hybrid subscription-OA model is much more likely to endure. During this more optimistic phase, I argued that grey literature advocates could learn from the political strength of open access advocates, and mount a similar campaign to demonstrate the value of grey literature* ^[5]. *Grey literature is almost always free to read already, so it only needed to be found.*

Today I am much less concerned about findability for grey literature. While it remains simpler to locate a journal article than a working paper, smart Google searches can easily unearth the latter (I'll provide some examples of this in the next section.) Portals to grey literature remain useful for providing context and browsability; along with the Grey Literature Report, OpenSIGLE also serves both these functions well ^[6]. *Even so, the core challenge of finding grey literature in the first place is much less potent than in years past. So rather than mounting a political campaign to raise the profile of grey literature, I believe that grey literature advocates should now concern themselves with strategies for preserving the ephemeral "grey data" represented in content such as blog posts and tweets* ^[7]. *After addressing improved findability, I will present the case for my position.*

Improved Findability for Grey Literature

Traditional barriers to locating grey literature, in comparison to white literature, include irregular publication schedules and the lack of standard bibliographic identifiers such as volumes, issues, and page numbers. These difficulties persist, but are much less fatal in the Internet age than they were in the print-only era.

It is possible to search Google for specific file types, and/or to restrict the search to the domains of organizations that produce a significant quantity of grey literature. Examples of each type of search, with screen shots, are found below (screen shots are current as of September 5, 2009.)

* First published in *Grey Literature in Library and Information Studies*. – Berlin : De Gruyter Saur, 2010.

Figure 1: Search for specific file type: "health policy filetype:pdf"

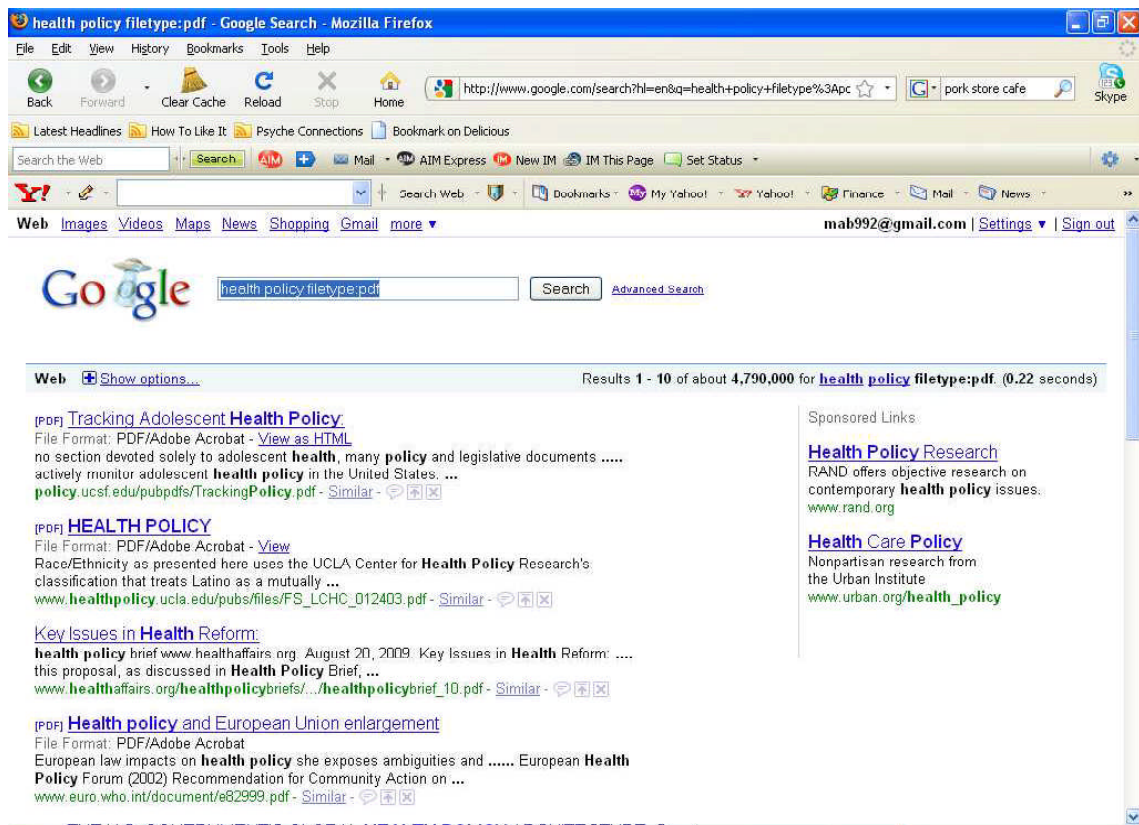
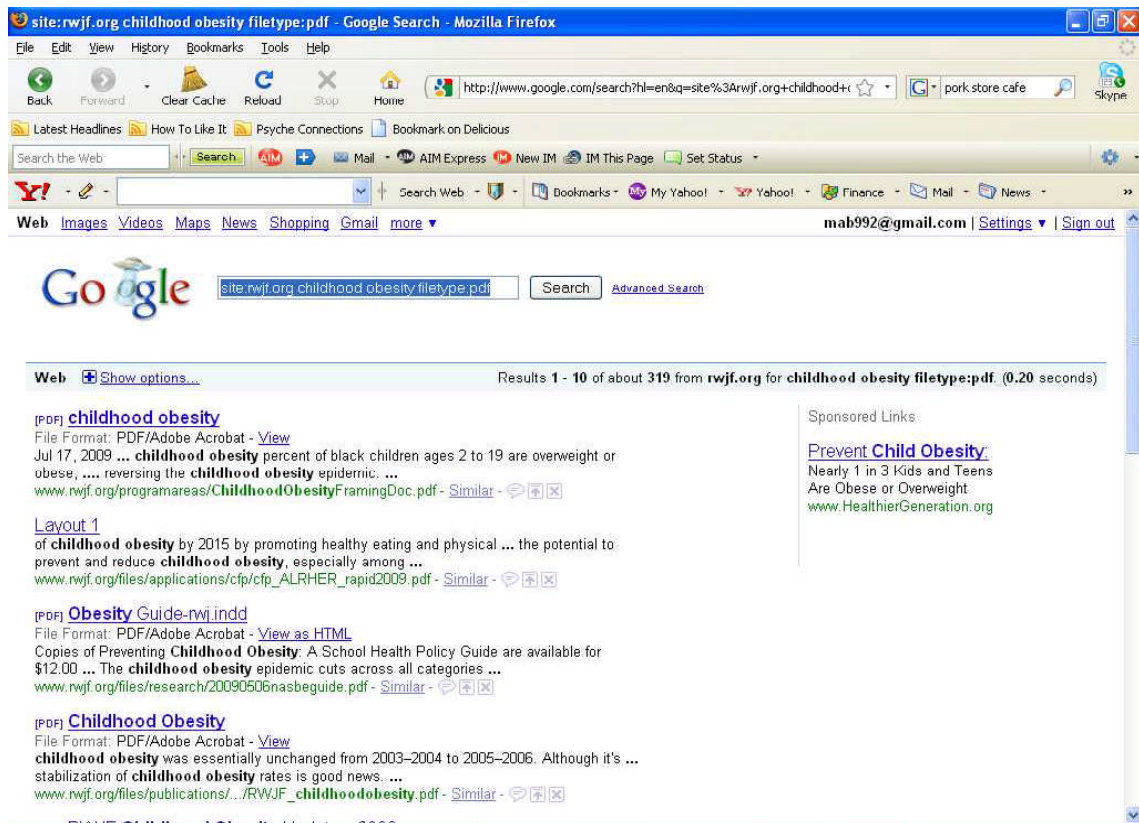


Figure 2: Search within domain of specific organization
site:rwjf.org childhood obesity filetype:pdf
(Restricts to searches of the Robert Wood Johnson Foundation domain)



These examples demonstrate how easy it is to locate grey literature, *as long as searchers know the appropriate search strings*. I recognize that this syntax is obscure to a very large majority of end users. However, librarians can instruct patrons about these search strategies, and also develop search systems that invoke these kinds of strategies behind the scenes.

Findability is not the only concern with grey literature. In the first example above the documents come from a large smattering of organizations, and would be of varying use depending on an individual's information needs. In the second search, I needed to know in advance that the Robert Wood Johnson Foundation was a useful source of information. The advantage of OpenSIGLE and the Grey Literature Report, in comparison to "raw Google," is that they both organize materials and vet their sources.

That said, findability *per se* is no longer a concern for traditional grey literature.

"Web 2.0" Content as "Grey Data"

If we accept the premise that findability is a less pressing concern than before, how should the grey literature community focus its energies today? I submit there is a pressing need to preserve the content being generated via various Web 2.0 tools and platforms.

First, some background: the somewhat annoying moniker "Web 2.0" describes the more interactive Internet that has emerged in recent years^[8]. First generation web sites tended to be static HTML pages where readers could look but not touch, and the only way to share anything was to forward web links via email. Today it is extremely easy to post articles or news clips (say, from the *New York Times* or the BBC) to one's Facebook page. (Facebook is just one of many "social networking" services; others include LinkedIn or MySpace.) This enhances sharing between friends and colleagues, which is useful. But the most profound change lies in the ability for anyone to post "user-generated content" such as blog posts or YouTube clips. Web 2.0 tools are also beginning to influence scientific debate^[9].

Blogs are now an established part of the information landscape; they are scrolling public diaries that usually allow comments. Twitter is a more recent development. It is a "micro-blogging" service that enables users to post very short messages (no longer than 140 characters) via their web browsers or mobile phones. Each message is a "tweet," and "retweeting" interesting messages has emerged as a rapid way to broadcast information^[10].

I have maintained a blog—which contains ruminations on both professional and personal matters—since January 2005^[11]. Many of my colleagues in health sciences libraries also write blogs, often with a mix of personal and professional content^[12, 13, 14]. In 2004 I did not read any blogs for professional information, and today they are critical professional sources. Blogs are updated much more rapidly than traditional journals, and readily facilitate conversation (although people often do not comment on posts.) Many librarians now tweet, sometimes to make pithy observations but often to share interesting blog posts or other online content^[15, 16]. Within Facebook, the status update is functionally similar to a tweet.

The various Web 2.0 tools have enhanced both my professional and personal life. They enable a more fluid and informal form of communication, and in some form they are here to stay. (This Web 2.0 stuff could all be a fad, but at least within the library realm I think it's unlikely that we'll return to journal articles and white papers as dominant distribution mechanisms.) If this premise is correct, then a major concern about reliance on Web 2.0 tools is that we do not yet have good mechanisms for permanently archiving content produced with these tools.

Tweets are almost by definition ephemeral, and blog posts suffer from the general "link rot" that bedevils the Web^[17]. Commercial solutions for Twitter archiving are emerging^[18], and at least we have the Internet Archive^[19]. But we are not yet close to an equivalent to acid-free paper for content developed online. Grey literature advocates can step into this breach, particularly if we expand the definition of grey literature to include the more informal "grey data"^[20].

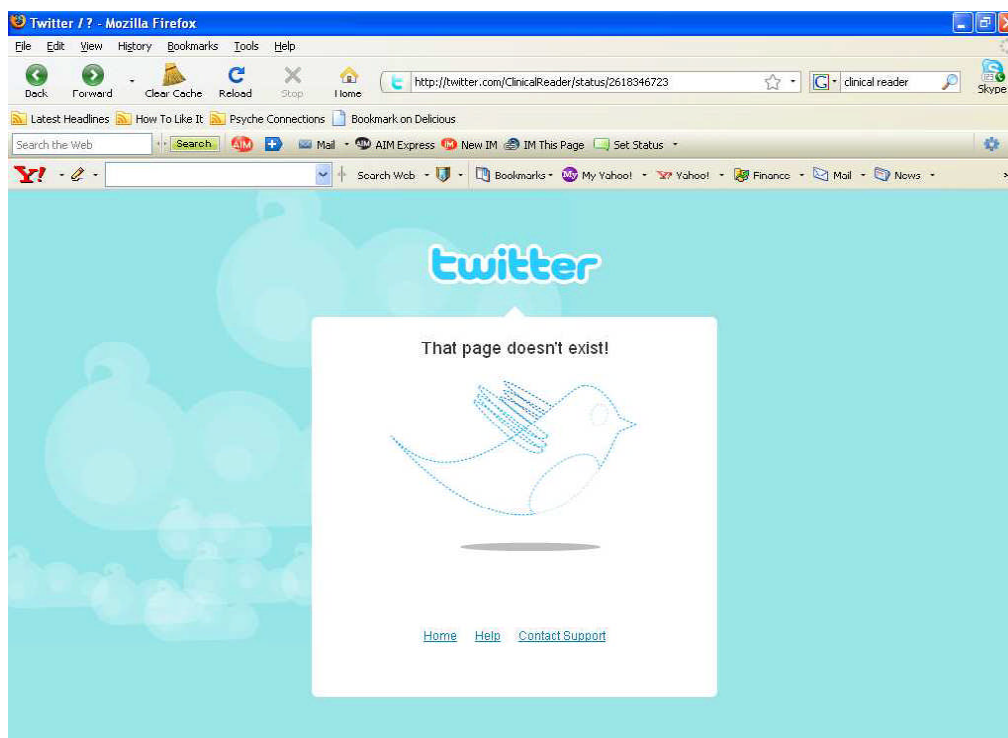
In the meantime, there are ad hoc efforts to preserve Web 2.0 content. For example, in the summer of 2009 the US Library of Congress announced plans to preserve tweets associated

with Justice Sonia Sotomayor's successful nomination to the Supreme Court ^[21]. An endorsement of the value of tweets (and by extension of blog posts, as many of the tweets referred to blogs commenting on the nomination) by an institution such as LC is a powerful indicator of their importance. Given LC's decision, to close this chapter I will provide further arguments for why the grey literature community should accept the challenge of preserving Web 2.0 content.

Case Study: Nicole Dettmar and Clinical Reader

Before making the argument for preservation of Web 2.0 content as grey literature or "data," let's examine the case of my librarian colleague Nicole Dettmar. Her experience in the summer of 2009 points to the vital need for permanent archiving of tweets in order to understand how people communicate online today.

Dettmar blogs, and in July 2009 she skillfully criticized the web site Clinical Reader for falsely implying that it had earned endorsements from leading libraries and for using copyrighted images without permission ^[22]. The critique was both accurate and thoughtful. Clinical Reader's initial response, via Twitter, was to "kindly request" that Dettmar remove her blog post or else face the risk of legal action ^[23]. As of September, Clinical Reader has apologized to Dettmar and removed the implication of non-existent endorsements.



In the wake of Clinical Reader's overwrought response to her post many librarians, as well as *Guardian* columnist Ben Goldacre, leapt to her defense. Much of the conversation between Clinical Reader and its critics took place on Twitter. But it was a hard conversation to have, because representatives from Clinical Reader consistently deleted their tweets and re-emerged with new Twitter accounts. Dettmar utilized screen shots to preserve the tweets for posterity ^[24], and the Disruptive Library Technology Jester also tracked the action ^[25]. Despite these laudable efforts, it would now be difficult for even the most intrepid scholar to piece together what happened on Twitter because of the deliberate interruptions to the flow of conversation. Above is a screen shot of a Clinical Reader Twitter screen that no longer exists, because the content was deliberately deleted.

If there is anyone who should care passionately about the preservation of otherwise overlooked and forgotten discourse, it is the grey literature community.

As Dettmar modestly stated, her initial post about Clinical Reader “was not of massive importance” ^[26]. But the support she received was gratifying and invigorating. Another important dimension of this support was its technical fluency. Dettmar’s supporters utilized Twitter with ease, responding to her critics within Twitter and retweeting in order to bring more attention to aspects of the discussion ^[27]. All of this happened in real-time, at a much faster pace of discourse than existed before the Web, or even in the Web 1.0 days. Tweets are early warning devices, and the blog posts or news articles they reference provide the context for whatever controversy is brewing.

Whether or not one has much interest in the Clinical Reader controversy, the phenomenon of the “rapid stream” of comments that registered the controversy should be of interest to the grey literature community. As I write, United States residents and political leaders are debating whether and how to reform the health care system. Here too Twitter offers a vital register of the discussion. And just as with the Clinical Reader tweets, there is no guarantee that this record will persist. Unless, that is, members of the grey literature and broader information communities resolve to preserve this record and others like it.

Connection Between Web 2.0 Content and Grey Literature/Grey Data

My initial interests in grey literature stemmed from admiration at the political savvy of open access advocates, with hopes that the GL community could learn from the open access movement and raise the profile of grey literature ^[28]. Now that findability for grey lit is not as large a concern, my interests have shifted to the idea of a “continuum” that will eventually collapse the distinction between grey and white literature ^[29].

I once assumed that this continuum only included materials that could easily be printed. My idea was that peer review could often happen online rather than behind the scenes, and I conceived of blog posts as the most radical extension of traditional forms of communication. Everything I envisioned would properly fall under the heading of “literature.”

As tweets can be no longer than 140 characters, it is a stretch to call them literature. But they are definitely useful bits of data, which collectively can aggregate into an important lens of understanding an ongoing discussion. In 2009 the lonely tweet suffers from the same findability problems as the working paper did pre-Google. Twitter facilitates searching for groups of tweets via the “hash tag” convention (example: #Obama), but the individual tweets buried within conversations can be easily lost ^[30].

If we conceive of the tweets as “grey data,” then the preservation imperative for the grey literature community becomes clearer. Robust archiving services for Twitter would have prevented Clinical Reader from manipulating the Twitter conversation this year.

Blog posts leave a larger footprint than tweets, although they still ephemeral and subject to the usual Internet link rot. They are also “grey” in the sense of not containing standard bibliographic identifiers. Although tweets are particularly at risk because they are so easy to proliferate, blog posts would also be worthy of the preservation attention of the grey literature community.

This begs a whole host of questions. With literally millions of bloggers of Twitterers out there, how can anyone possibly determine what to preserve? The Library of Congress’s decision regarding the Sotomayor tweets was an easy call, but this won’t always be the case. While I recognize the enormity of the challenge, the developers of the New York Academy of Medicine and OpenSIGLE portals also had to establish selection criteria (albeit on a much smaller scale). General digital preservation principles, combined with an evolving understanding of the uses of Twitter, would be necessary in developing preservation criteria for blogs and tweets. I have no words of wisdom in this regard, except to say that this is clearly an area of growth for the grey literature community and that I would be happy to be part of any discussions in these areas.

Conclusion

My interests in grey literature have shifted considerably in the last five years, which is roughly synonymous with the emergence of “Web 2.0” tools. I almost feel like a fraud writing this chapter, because my interests are so divergent from what I presented at conferences in Nancy and New Orleans! But at the risk of seeming fraudulent, I really do believe that preservation of Web 2.0 content should be a main focus for the GL community in future years. Traditional grey literature remains important, but thankfully it is much easier to find than before. Let us now turn our attention to new and exciting challenges.

References

1. Alpi K.M. (2005), Expert searching in public health. – In: J Med Libr Assoc, 93 (1) Jan 2005, pp. 97-103. <http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=545128&blobtype=pdf>
2. NLM, Partners in Information Access for the Public Health Workforce, <http://phpartners.org/>. [cited 7 Sep 2009].
3. NYAM Library, Grey Literature Report, <http://www.nyam.org/library/greyreport.shtml>. [cited 7 Sep 2009].
4. Suber P. (2003), Bethesda statement on open access publishing, <http://www.earlham.edu/~peters/fos/bethesda.htm>. [cited 7 Sep 2009].
5. Banks M. (2004), Connections between open access publishing and access to gray literature. – In: J Med Libr Assoc, 92 (2) April 2004, pp. 164-166. <http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=385294&action=stream&blobtype=pdf> [cited 7 Sep 2009].
6. INIST, OpenSIGLE, <http://opensigle.inist.fr/>. [cited 7 Sep 2009].
7. Banks M. (2009), Open access, grey literature, grey data, http://mbanks.typepad.com/my_weblog/2009/07/open-access-grey-literature-grey-data.html [cited 7 Sep 2009].
8. O'Reilly T. (2005), What is Web 2.0?, <http://oreilly.com/web2/archive/what-is-web-20.html>. [cited 7 Sep 2009].
9. Economist (2008), User-generated science, http://www.economist.com/sciencetechnology/displayStory.cfm?STORY_ID=12253189 [cited 7 Sep 2009].
10. Blogging Bits (2008), How to retweet: a simple guide, <http://bloggingbits.com/the-art-and-science-of-retweeting-for-twitteraholics/>. [cited 7 Sep 2009].
11. Banks M., Marcus' World, <http://mbanks.typepad.com/> [cited 7 Sep 2009].
12. Kraft M., Krafty Librarian, <http://kraftylibrarian.com/> [cited 7 Sep 2009].
13. Plutchak T.S., T. Scott, <http://tscott.typepad.com/>. [cited 7 Sep 2009].
14. Rothman D., Exploring medical librarianship and web geekery, <http://davidrothman.net/>. [cited 7 Sep 2009].
15. Rumsey E., Eric Rumsey on Twitter, <http://twitter.com/ericrumsey>. [cited 7 Sep 2009].
16. Walden R., Rachel Walden on Twitter, http://twitter.com/Rachel_w. [cited 7 Sep 2009].
17. Wikipedia (2009), Link rot, http://en.wikipedia.org/wiki/Link_rot. [cited 7 Sep 2009].
18. QuoteURL (2009), <http://www.quoteurl.com/>. [cited 7 Sep 2009].
19. Internet Archive (2009), <http://www.archive.org/index.php>. [cited 7 Sep 2009].
20. Banks M. (2009), Open access, grey literature, grey data, http://mbanks.typepad.com/my_weblog/2009/07/open-access-grey-literature-grey-data.html. [cited 7 Sep 2009].
21. Library of Congress (2009), LOC will capture tweets on #sotomayor for its web archives on the Sotomayor nomination, <http://twitter.com/librarycongress/statuses/2222610835>. [cited 7 Sep 2009].
22. Dettmar N. (2009), Clinical Reader: Starry ethics fail, <http://eagledawg.blogspot.com/2009/07/clinical-reader-starry-ethics-fail.html> [cited 7 Sep 2009].
23. Lawson S. (2009), Clinical Reader: From zero to negative sixty with one bogus threat, http://stevelawson.name/seealso/archives/2009/07/clinical_reader_from_zero_to_negative_sixty_with_one_bogus_threat.html [cited 7 Sep 2009].
24. Dettmar N. (2009), Twitter search on “legal ramifications,” July 16, '09 on Twitpic, <http://www.twitpic.com/al4eg> [cited 7 Sep 2009].
25. Disruptive Library Technology Jester (2009), On the pitfalls of social media: The case of Clinical Reader, <http://dlitj.org/article/clinical-reader-background/> [cited 7 Sep 2009].
26. Dettmar N. (2009), Gratitude, <http://eagledawg.blogspot.com/2009/07/gratitude.html> [cited 7 Sep 2009].
27. Lozeau P.M. (2009), RT @adr Dear @ClinicalReader: Stop threatening bloggers with bogus legal claims. OK? Thanks. (i.e <http://ur1.ca/78oc>). [cited 7 Sep 2009].
28. Banks M. (2004), Connections between open access publishing and access to gray literature. – In: J Med Libr Assoc, 92 (2) April 2004, pp. 164-166. <http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=385294&action=stream&blobtype=pdf> [cited 7 Sep 2009].
29. Banks M. A. (2005), Towards a continuum of scholarship: the eventual collapse of the distinction between grey and non-grey literature. – In: GL7 : Seventh International Conference on Grey Literature, Nancy (France), December 2005. <http://eprints.rclis.org/5803/> [cited 7 Sep 2009].
30. Tech for Luddites (2009), The Twitter hash tag: what is it and how do you use it?, <http://www.techforluddites.com/2009/02/the-twitter-hash-tag-what-is-it-and-how-do-you-use-it.html> [cited 7 Sep 2009].

Fourteenth International Conference on Grey Literature



Tracking Innovation through Grey Literature

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

Hosted by the Central Library "Guglielmo Marconi", IRPPS Rome, and ISTI Pisa

Announcement

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

The goal of the Fourteenth International Conference on Grey Literature seeks to track the process of innovation by tracing the research life cycle and observing its convergence in the field of grey literature. The recommended conference topics and their interrelated threads as appear below provide ready and ample inroads for the GL14 Call-for-Papers.

Conference Topics and Treads

New Document Types and Collections	... Electronic Theses and Dissertations (ETDs), Blogs, Audio-Visuals, Databases, Software, and Post-publication Data
Tracing the Research Life Cycle	... Research Data, Leading edge, First-in, Evidence based, Research Results
Tracking Methods	... Publication Life Cycle, Citation Data, Persistent Identifiers, Weblogs, User Stats, References, Impact Factors
Adapting New Technologies	... Standards for Data Formats, New Search and Communication Tools, Managing the Information Workflow and Review Process
Repurposing Grey Literature	... Sustainability, Public Accessible Resources, Product and Service enhancements, Open Access, Curation and Preservation
Linking Research and Teaching	... Knowledge Transfer, Policy Implications, Decision Making, Problem Solving, Good Practice, Cross-Curricula Instruction, Raising Awareness and Profile
New Image of Librarians	... Expanding the Library Landscape and its Resources, Involvement of other Information Professionals, Practitioners, and Stakeholders
Community and Crowd Sourcing	... Social Networking, Community Driven not Mandated, Focus on the Supply and Demand Sides, Targeted Audiences as well as Wider Public

Note In order to keep the conference delegates, authors, and participants informed on the progress of the Fourteenth International Conference on Grey Literature, a quarterly newsletter, conference calendar, and further documentation will be maintained on the conference site.

Fourteenth International Conference on Grey Literature



Tracking Innovation through Grey Literature

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

Hosted by the Central Library "Guglielmo Marconi", IRPPS Rome, and ISTI Pisa

Call for Papers

Guidelines for Abstracts

Title of Paper:

Author(s) Name(s):

Organization(s):

Address/P.O. Box:

Postal Code - City – Country:

Topic(s) as shown below:

Telephone:

Fax:

Email:

URL:

Participants who seek to present a paper at GL14 are invited to submit an English abstract between 300-350 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, 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 will be returned to the author for revision.

Conference Topics and Trends

New Document Types and Collections	... Electronic Theses and Dissertations (ETDs), Blogs, Audio-Visuals, Databases, Software, and Post-publication Data
Tracing the Research Life Cycle	... Research Data, Leading edge, First-in, Evidence based, Research Results
Tracking Methods	... Publication Life Cycle, Citation Data, Persistent Identifiers, Weblogs, User Stats, References, Impact Factors
Adapting New Technologies	... Standards for Data Formats, New Search and Communication Tools, Managing the Information Workflow and Review Process
Repurposing Grey Literature	... Sustainability, Public Accessible Resources, Product and Service enhancements, Open Access, Curation and Preservation
Linking Research and Teaching	... Knowledge Transfer, Policy Implications, Decision Making, Problem Solving, Good Practice, Cross-Curricula Instruction, Raising Awareness and Profile
New Image of Librarians	... Expanding the Library Landscape and its Resources, Involvement of other Information Professionals, Practitioners, and Stakeholders
Community and Crowd Sourcing	... Social Networking, Community Driven not Mandated, Focus on the Supply and Demand Sides, Targeted Audiences as well as Wider Public

Due Date and Format for Submission

Abstracts in MS Word must be emailed to conference@textrelease.com by April 12th 2012. The author will receive verification upon receipt. Shortly after the Program Committee meets on May 4th 2012, 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, PowerPoint slides, and Conference Registration.

Banks, Marcus

Marcus A. Banks is the director of Samuel Merritt University's division of Library/Academic & Instructional Innovation. He was previously the Manager of Education and Reference Services at the University of California San Francisco. Marcus is also the recipient of the GreyNet Award 2006. He remains interested in how the concept of grey literature will evolve in the digital age. Email: mbanks@samuelmerritt.edu

Davis, Marilyn J.

Lynn Davis is an information scientist and account executive for the U.S. Department of Energy's (DOE) Office of Scientific and Technical Information (OSTI) in Oak Ridge, Tennessee. She serves DOE, other government agencies, and other organizations by developing solutions to provide access to DOE scientific and technical information meeting their specific needs. Examples of her successful solutions and collaborations include managing an historic legacy digitization project for the U.S. DOE Geothermal Technologies Program (osti.gov/geothermal); developing new media services for the DOE National Science Bowl to lay and build upon a foundation for fostering engagement among students, teacher, parents, community, academic sponsors, and the media around an annual nationwide academic competition that tests students' knowledge in all areas of science (osti.gov/nsbjournal); and leveraging the abundance of scientist/researcher contact information in OSTI databases to facilitate peer-to-peer communication, collaboration, and sharing between and among scientists (www.Science-Forums.net). Ms. Davis has a Bachelor of Science from James Madison University and a Master of Science in Information Sciences from the University of Tennessee. Email: DavisL@osti.gov

Gelfand, Julia

Julia Gelfand has been engaged with Grey Literature for nearly 20 years and has written and spoken extensively about the role of grey literature as it aligns with contemporary publishing and its role in library collections. She has explored the new fringe of grey literature in virtual worlds, evidence-based practices and now in social media and is interested in how it will be tracked, preserved and archived. She is the Applied Sciences & Engineering Librarian at the University of California, Irvine where she has been a member of the staff for over 30 years and is actively involved in collection development, scholarly communication, and e-science activities at national and international levels. Email: jgelfand@uci.edu

Južnič, Primož

Primož Južnič is an associate professor at the Department of Library and Information Science and Book Studies at Faculty of Arts, University of Ljubljana (Slovenia). His main area of research and interest is bibliometrics, collection management and LIS education. He teaches the following courses: Bibliometrics, Special libraries, and Collection Management. Before starting his university career, he was a heading different special and academic libraries and information/computer centres. He was also working at the European Commission, for three years, as the seconded informatics expert. Email: primoz.juznic@ff.uni-lj.si

Lin, Anthony

Tony Lin is the Research Librarian for Business at the University of California, Irvine. He holds a MIS from the University of Michigan-Ann Arbor, School of Information. At the University of California, Irvine, he utilizes and actively explores new ways to use emerging library technologies to effectively deliver reference services to both on-campus and distance education users. In addition to his work with emerging technologies, he also works closely with small business owners and entrepreneurs to form business plans, perform competitive analysis, and market research for local and regional businesses. His many interests include Geographic Information Systems, US Government Information, and institutional data repositories. Email: ablin@uci.edu

Lin, Yongtao

Yongtao Lin has been a Health Information Network librarian, working at the Tom Baker Cancer Knowledge Centre since 2008. She provides library services to support health care professionals in their evidence-based practice. The library is also part of a provincial patient-centered education strategy that supports cancer patients and their families. She was a librarian with hospital libraries in rural Nova Scotia for a few years before working with the University of Calgary. Her prior experience being an instructor has led her to integrate education into various aspects of library programs. Yongtao is interested in the impact of grey literature in health care and a strong believer of evidence-based practice. yolin@ucalgary.ca

Marzi, Claudia

Claudia Marzi Position: Researcher (CNR Tecnologo) PhD student in Acquisitional and Computational Linguistics, University of Pavia. Laurea degree in Modern Languages (English) at Pisa University in 1998, with the dissertation "The power of words: language creativity in Edgar Allan Poe's narrative". Since 2001 working at the Institute for Computational Linguistics "A. Zampolli" (ILC), National Research Council (CNR), in Pisa. She contributed to many national (funded by Italian Ministry of Research) and international (funded by European Commission) projects. Programme Coordinator of the European Science Foundation Research Networking Programme "The European Network on Word Structure. Cross-disciplinary approaches to understanding word structure in the languages of Europe". Member of board at Institute for Computational Linguistics. Main areas of interest: Document and knowledge management; Second language acquisition; Child language; Computer modelling of the Mental Lexicon. Email: claudia.marzi@ilc.cnr.it

Ranjit, Kathryn M.E.

Kathryn M.E. Ranjit is a librarian at the Peter Loughheed Knowledge Centre, Health Information Network. She received her MLIS and BMSc from the University of Western Ontario. She enjoys the challenges of providing quality information service to health care professionals, patients and families and seeks to implement evidence-based decision making in her practice. Her current areas of research include web development to facilitate the online access to health information and the use of social media for online reference. She is also an advocate for the impact of health literacy on health care. Email: kathryn.ranjit@ucalgary.ca

Vowell, Lance

Lance Vowell is currently employed by Information International Associates and assigned to the U.S. Department of Energy's (DOE) Office of Scientific and Technical Information (OSTI) contract as the Contract/Program Manager. In this capacity, Mr. Vowell is responsible for all phases of IT support ranging from requirements gathering, deployment and product maintenance. Mr. Vowell has been instrumental in the development of core application framework that has markedly reduced application development time and maintenance costs at OSTI. Vowell has served as a technical resource and/or Principal Investigator on six Small Business Innovation Research (SBIR) grants for the DOE on topics including "Contextual Spell Checking in the Digital Library Environment", "Mobile Device Applications in the Digital Library Environment" and "Interactive Peer to Peer Scientific Communication in the Digital Library Environment"; two Small Business Technology Transfer (STTR) grants for DOE and the U.S. Department of Defense; and is currently serving as a technical advisor on an DOE Experimental Program to Stimulate Competitive Research (EPSCoR) grant, "Exploiting the Use of Social Networking to Facilitate Collaboration in the Scientific Community". VowellL@osti.gov

Notes for Contributors

Non-Exclusive Rights Agreement

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

Submission Requirements

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

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

REFERENCE GUIDELINES

General

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

Specific

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

Examples

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

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

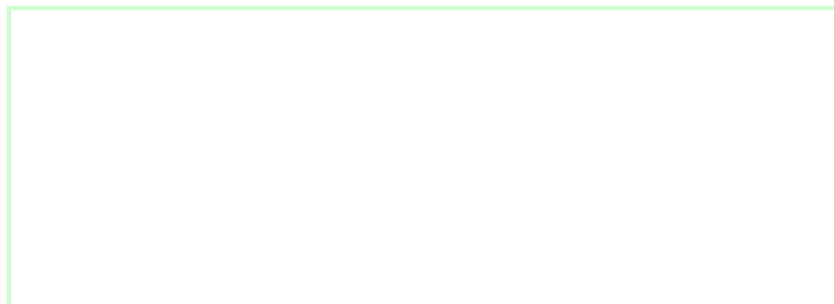
DCMI, Dublin Core Metadata Initiative Home Page http://purl.oclc.org/metadata/dublin_core/

Review Process

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

Journal Publication and Article Deposit

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



An International Journal on Grey Literature

'SOCIAL NETWORKING AND GREY LITERATURE'

SPRING 2012 – TGJ VOLUME 8, NUMBER 1

Science-Forums.net: A Platform for Scientific Sharing and Collaboration	5
<i>Marilyn J. Davis, Edrick Coppock, and Lance Vowell (United States)</i>	
Social Networking: Product or Process and What Shade of Grey?	14
<i>Julia Gelfand and Anthony Lin (United States)</i>	
Knowledge Communities in Grey	27
<i>Claudia Marzi (Italy)</i>	
Using Social Media to Create Virtual Interest Groups in Hospital Libraries	35
<i>Yongtao Lin and Kathryn M.E. Ranjit (Canada)</i>	
Information-Seeking Behaviour of Slovenian Researchers: Implications for information services	43
<i>Polona Vilar, Primož Južnič, and Tomaz Bartol (Slovenia)</i>	
Blog Posts and Tweets: The Next Frontier for Grey Literature	54
<i>Marcus Banks (United States)</i>	

Colophon	2
Editor's Note	4
Advertisement	
New York Academy of Medicine	34
GL14 Conference Announcement	60
GL14 Call for Papers: Tracking Innovation through Grey Literature	61
About the Authors	66
Notes for Contributors	67