e-ISSN 1574-180X

An International Journal on Grey Literature

Volume 16, Number 2, Summer 2020

'GREY LITERATURE – A PLATFORM AND INTERFACE FOR OPEN SCIENCE'

GreyNet
www.textrelease.com Grey Literature Network Service www.greynet.org

The Grey Journal

An International Journal on Grey Literature

COLOPHON

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.

CIP
The Grey Journal (TGJ) : An international journal on grey literature / Dominic Farace (Journal Editor); Jerry Frantzen (Technical Editor) ; GreyNet International, Grey Literature Network Service. - Amsterdam: TextRelease, Volume 16, Number 2, Summer 2020 – TIB (DE), DANS-KNAW (NL), CVTISR (SK), EBSCO (USA), ISTI CNR (IT), KISTI (KR), NIS IAEA (UN), NTK (CZ), and the University of Florida (USA) are Corporate Authors and Associate Members of GreyNet International.
e-ISSN 1574-180X (PDF)

Subscription Rate:
€240 institutional

Contact Address:
Back Issues, Document Delivery, Advertising, and Subscriptions:

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

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.

Contents

'Grey Literature – A Platform and Interface for Open Science'

EDITOR'S NOTE

Data Papers are an Innovative Tool for Information and Data Management

In earlier work dealing with Data Papers, emphasis initially focused on its definition, the construction of a standardized template in compiling its component parts, its further publication, access and data preservation. This was followed by a study seeking to engage the citation and referencing of data papers and the further sharing and reuse of the data described therein. A more recent work outside GreyNet's immediate community of practice addressed the diverse formats used to compile and publish data papers. That study also discussed the automatic generation of data papers, and the differing opinions as to whether they contribute to data, information and/or knowledge production.

Drawing on the results of the above-mentioned work, a new study seeks to demonstrate how the data paper provides an innovative tool for information and data management, as part of an "ecosystem" of conference proceedings, journal articles, research data and open repositories. It relies upon GreyNet's current collection of 40 plus published datasets and 12 plus data papers – two of which appear published in this issue of The Grey Journal.

This study highlights the importance of the human contribution for the writing of data papers and the enrichment of their metadata. To this end, key shared components of GreyNet's collection of data papers are discussed, namely the stakeholders, linked metadata, open data archiving, preservation, and issues of quality and information rights.

Dominic Farace,
Journal Editor

Is the Production and Use of Grey Marine Literature a Model for Open Science?*

Bertrum H. MacDonald, Rachael Cadman, Curtis Martin, Simon Ryder-Burbidge, Suzuette S. Soomai, Ian Stewart, and **Peter G. Wells;** Dalhousie University, Canada

Abstract

Globally, grey literature is common. Large quantities of openly available grey literature have been generated since the latter half of the nineteenth century. It is a primary source of information used in many public policy and decision-making contexts, at all jurisdictional levels. In fact, public decision making and policy development would seriously falter today in the absence of such literature. Moreover, in some jurisdictions, legislation mandates transparent governance processes in which current research must be fully open. This lengthy experience with open practices in the production and use of grey literature offers insights to the open science movement.

In this paper, based on over fifteen years of interdisciplinary research, we demonstrate how open practices in the production and use of grey literature in marine environment science policy contexts could inform open science initiatives. The results from our numerous case studies about information use in decision-making processes, at local to global levels, address two conference themes, namely, the application of open science principles in promoting grey literature, and obstacles and challenges to such open access.

Information pathways in coastal and ocean management are complex and involve many actors (including researchers; managers; policy analysts; members of industry, professional associations, community groups, and non-governmental organizations; politicians; and citizens generally). Open grey literature offers numerous advantages in these settings, as an extensive variety of information needs, types, and formats are prevalent. Open grey literature can also be distributed without restriction by individuals and organizations. It can now be shared globally with ease, which is particularly beneficial to developing countries often unable to afford commercial information sources.

However, while produced and used widely, grey literature also presents challenges that open science also encounters. Openness, i.e., open access, does not ensure awareness and it does not automatically equate to usability by a wide variety of audiences. Because grey literature is assumed to be largely accessible, often limited attention is focused on promoting awareness or communicating information in broadly understandable terms. Furthermore, the massive quantity of literature can contribute to its seeming invisibility. The multiplicity of formats and content can result in perceptions of limited value of grey literature. Even though the information may be rigorously peer-reviewed, in today's information-saturated environment, open-access may be equated with uncertain quality.

Our research on the use and influence of grey literature in marine environmental decision making highlights the benefits and challenges of open access information. Thus, our findings may be particularly informative to current efforts to advance open science principles globally.

We live in a paradoxical information world. To say that we are awash with information is a substantial understatement today. Everyone, from private citizens to public decision makers, regularly encounters an overabundance of information and must implement strategies and filters to select or deflect incoming information (Walgrave & Dejaeghere, 2017). Barriers are frequently established to mitigate the deluge of information. In contrast, advocates of open science have recently been promoting *greater* access to information (Open science, n.d.; Open science, 2019; Vincente-Saez & Martinez-Fuentes, 2018). As the name implies, open science initiatives emphasize unrestricted availability to the data and information emanating from research activity. For several years, leading science journals, e.g., *Science,* have required authors to submit the raw data generated by their research, for online publication as a supplement to research papers. Granting

* First published in the GL21 Conference Proceedings, February 2020.

councils often require the same practice in their reporting requirements. Impediments preventing access are being removed by publishers today or are not set in place at all in new research work. In short, coping mechanisms are established to limit or obstruct information flow on the one hand, while on the other hand, mechanisms are being implemented to facilitate access to the original data.

As matters of concern, information overload and open science are not new phenomena. Both have existed in some form since at least the fifteenth century (Blair, 2010; Muffler-Wille & Charmantier, 2016; Ogilvie, 2003; Rosenberg, 2003; Tidline, 2002). However, the attention both are receiving currently is a reflection of the very rapid growth in the volume of data and information over the last half-century and major advances in widely available technologies that facilitate production of scientific data, e.g., faster computers, automated instruments, satellites, as well as widespread distribution and access to information (Allen & Mehler, 2019; Landhuis, 2016).

The history of science demonstrates that scientists and their professional associations have long been champions of wider access to research data and information, as well as experiencing at times the opposite pressures to control and manage access for complex reasons (Vermeir & Margócsy, 2012). The recent literature on the subject of open science illustrates numerous aspects on this topic: open access, open data, open reproducible research, open science evaluation, open science policies, open science tools, among others (e.g., Méndez, 2019; Open science definition, n.d.). However, a consensus about the definition of open science has not yet coalesced around any one of these terms. The organizers of this conference adopted the definition of open science as "the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society" (following the explanation by Woelfle, Olliaro, and Todd (2011), now adopted by Wikipedia (Open science, 2019)). A recent systematic review of literature considered 75 studies on the subject and proposed the following definition: "Open science is transparent and accessible knowledge that is shared and developed through collaborative networks" (Vincente-Saez & Martinez-Fuentes, 2018). This succinct characterization of open science highlights two facets, collaborative knowledge development and wide accessibility of the results, whereas the previous definition involves widespread behavioural change that supports accessibility. Open science should consider all three descriptors in its definition.

In our collaborative research on questions about information production and its use at the science-policy interface in marine environmental and fisheries contexts (www.eiui.ca), we have observed and documented the extensive use of grey literature in decision-making processes. Collaboration in knowledge generation and numerous efforts to present information in meaningful ways to diverse audiences is evident in our findings. Large depositories of openly available marine data and information, published in various forms of grey literature, are generated specifically to support the operations of governments at all jurisdictional levels (MacDonald et al., 2013).

Today, advances in information technologies are contributing to a blurring of distinctions between some forms of grey literature and primary research literature. The publication and distribution of both now often follow similar practices. Our research has shown that the production and use of grey literature in marine environmental science-based decision-making contexts can inform open science initiatives. The results from our numerous case studies, at local to global levels, address two conference themes: the application of open science principles in promoting grey literature, and the obstacles impeding open access. Hence, this paper has two objectives: to illustrate how research about grey literature can provide insights for the open science movement, and to suggest that a sharp distinction between grey and primary literature is becoming less relevant as open science develops in theory and practice.

The Environmental Information: Use and Influence Research Program

For almost two centuries, governments in many nations have established research bodies to conduct scientific research, whether internal with a national focus or externally in collaboration with other countries (MacDonald & Soomai, 2019; Oppenheimer et al., 2019). Created for a variety

of reasons (economic, cultural, environmental, etc.), the large number of research bodies found around the world have produced a "spectrum of types of research output, including print and digital formats, largely due to the diversity of audiences that governments aim to reach" (MacDonald & Soomai, 2019, p. 29). Substantial financial resources have been required to generate this volume of research literature and related publications emanating from the governmental and intergovernmental organizations. As a measure of accountability alone, governments often wish to determine whether various types of data and information have been used and to what impact (Wells, 2003).

The initial research conducted by the interdisciplinary Environmental Information: Use and Influence research program (www.eiui.ca), in which we work at Dalhousie University, was prompted by the question of accountability: were major international reports on marine pollution and the state of the oceans ever used? That first study (Cordes, 2004; MacDonald, Cordes, & Wells, 2004), which confirmed that the reports in question were widely cited, took us into explorations of the sphere of the science-policy interface and numerous case studies of information production and use by local and national governmental bodies, international intergovernmental organizations, and environmental non-governmental organizations. Our research findings have been presented at earlier iterations of this grey literature conference, to many other conferences in North American and Europe, and in oral briefings to governmental and non-governmental organizations. In addition, we have published our research results in marine policy and ocean management journals, information studies journals such as *The Grey Journal*, and in technical reports and a recent monograph (see www.eiui.ca; MacDonald et al., 2016).

Our research has shown that information pathways in ocean and coastal management are complex and involve many actors (including researchers; managers; policy analysts; members of various industries, professional associations, community groups, and non-governmental organizations; politicians; and citizens generally). Grey literature, much of which is designed for open access, is widely deployed for information transfer and communication in marine and ocean governance settings. The flexibility of grey literature publication options responds well to an extensive variety of information needs, witnessed by the numerous types and formats that are commonly prevalent. Grey literature is usually distributed without restriction by individuals and organizations and can now be shared globally with relative ease in the internet age. This development is particularly beneficial to developing countries often unable to afford commercial information sources. In our recent book about different dimensions of information use in integrated coastal and ocean management, we highlighted the significant roles that information published as grey literature fulfills in current management practices (MacDonald, Soomai, De Santo, & Wells, 2016).

Grey Literature as a Model for Open Science Practices

The production and use of grey literature, by many actors in marine management decision making over past decades, illustrates six characteristics that advocates are calling for in open science (Table 1). Research that draws on experience with grey literature can inform the application and promotion of these characteristics in open science initiatives.

Table 1. Grey Literature as a Model for Open Science

Features of Grey Literature Relevant to Open Science
Usability of the Information
Thorough Peer Review
Designed for Diverse Audiences
Synthesized Reports for Public Policy Contexts
State of the Environment Reports
Fisheries Scientific Reports
Reports and Documents Prepared by Environmental NGOs
Environmental Assessments and Impact Assessments
Data Published as Open Access
Accessible to Public Users

1. Usability of the Information

In 2002, a group of researchers in the United States published a report that identified three key features of information that ensures its usability, namely salience (or relevance), credibility, and legitimacy (Cash et al., 2002). Grey literature is used in many marine management decision-making processes because it displays these key characteristics of useful information. We know that governmental organizations at national, regional, and international levels rely heavily on using their own publications in decision making. One of us observed this use during a career with the Canadian government as a marine environmental scientist (Wells, 2016). This phenomenon was also seen during a study of major prolific producers of fisheries information: Fisheries and Ocean Canada (DFO), the national ocean governance authority; Northwest Atlantic Fisheries Organization (NAFO), a regional fisheries management body; and the Food and Agriculture Organization of the United Nations (FAO) (Soomai, 2017a; 2017b).

DFO in Canada and NAFO at the international level have well-defined processes for producing, communicating, and using information. Scientific information produced by scientists, often commissioned for specific decision-making purposes, is the primary source of advice. The fisheries science publications (grey literature) produced by employees of these organizations are preferred in policy-making as the information presented in this manner is timely, i.e., the publications are produced in annual cycles, and is relevant because it is in direct response to the fisheries management needs and questions. The scientific advice is prepared in response to specific managerial problems, and undergoes rigorous internal peer review, thereby resulting in the production of credible information available for immediate use in fisheries decision making. The decision processes also involve external stakeholders, e.g., the fishing industry, NGOs, and academic researchers, which reinforces the legitimacy and role of the information in the broader ocean governance process.

2. Peer Review

Peer review is widely viewed as an essential quality control mechanism. Peer review methods have been evolving over the past half century, and although none are entirely fool-proof, and some are contested (Bohannon, 2013; Haider, & Åström, 2017), peer review practices are considered essential in publishing credible research results (e.g., Baldwin, 2018; Lee & Moher, 2017).

The credibility of many research reports released as grey literature depends in large measure on the application of peer review. For example, for 50 years, the Joint Group of Experts on Scientific Aspects of Marine Environmental Protection (GESAMP), a leading international scientific advisory body, has been publishing substantial reports about the condition of the oceans (MacDonald, Cordes, & Wells, 2004; Wells, Duce, & Huber, 2002). GESAMP has published over 100 reports, including major periodic assessments, such as *The State of the Marine Environment* (No. 39) (GESAMP, 1990) and *Protecting the Oceans from Land-Based Activities* (no. 71) (GESAMP 2001). Each report is drafted in one of the working groups. Report No. 64 (GESAMP, 2002), for example, was the result of six years of work by the thirteen member Working Group on the Evaluation of the

Hazards of Harmful Substances Carried by Ships. Each draft report is subjected to extensive review, by both external reviewers and detailed, page-by-page consideration by GESAMP members in the annual sessions. Report No. 71, for example, lists over 90 individuals who had various roles in its preparation (GESAMP, 2001). The extensive and rigorous treatment of GESAMP's reports is an example of the degree to which this grey marine literature is peer reviewed before publication. Given the challenges of peer review sometimes faced by the move toward open science, e.g., prevalence of predatory journals with flawed to non-existent peer review (Bohannon, 2013), the review process practised by GESAMP could be an example for addressing this issue.

3. Designed for Diverse Audiences

In part because grey literature is not constrained by particular styles, formats, or communication channels, this literature offers considerable flexibility for creators to design information products for diverse audiences. Flexibility is important when decision processes encompass many stakeholder groups and end users. For example, our case study of the publications of the Gulf of Maine Council on the Marine Environment (GOMC), an international, intergovernmental organization involving two Canadian provinces, three American states, and the national governments of both countries, highlighted a diversity of information products. They include annual reports, brochures and posters, conference background papers, conference reports and workshop proceedings, fact sheets, technical reports, serials, journal articles, and abstracts (Cossarini, MacDonald, & Wells, 2014; MacDonald, Cordes, & Wells, 2007). This range of publications formats reflects the breadth of projects that the GOMC pursues and the span of audiences that it intends to reach in the multiple jurisdictions within its ambit. Flexibility in format is a benefit when a single project aims to engage researchers, environmental managers, policy makers, and the public. Such was also the case with the *State of Nova Scotia's Coasts Report*, which was released in two languages, print and digital formats, and as a 245-page technical report, a 21-page summary, and six four-page fact sheets. Publishing this variety was a deliberate decision to provide accessible and credible information to professional, government and lay audiences (Soomai, MacDonald, & Wells, 2013).

By designing information products for different stakeholders, whose capacity to understand scientific information varies, creators of grey literature can be successful in achieving transparency and accessibility to both professional and amateur audiences, both key elements in open science.

4. Synthesis Reports for Public Policy Contexts

Synthesis reports intended for use in public policy contexts are a common form of grey literature. The major synthesis reports produced by the Intergovernmental Panel on Climate Change (IPCC) are some of the most widely known, but its reports are only one of many such publications produced each year and freely accessible.

Numerous examples of synthesis reports can be drawn from our studies.

a) State of the environment reports
   i) *The 2009 State of Nova Scotia's Coast Report* (see Soomai, MacDonald, & Wells, 2013)
   ii) *The State of the Gulf of Maine Report* (see Soomai, MacDonald, & Wells, 2013)
   iii) The *State of the Scotian Shelf Report* (see Ross, 2014; Ross & Breeze, 2016)

   The rigorous and transparent methods by which these reports were prepared and distributed could inform open science efforts. For instance, a suite of methods (print and digital) was used to promote awareness and use of the respective reports and the methods were selected to reach various audiences (as described in section 3 above).

b) Fisheries scientific reports

   Fisheries management organizations (e.g., DFO, NAFO, and FAO) rely heavily on their own reports (grey literature) to inform national, regional, and international fisheries management decisions and policy making (as described in section 1 above) (Soomai, 2017a, 2017b).

c) Reports generated by advisory working groups or committees.

   Interdisciplinary working groups provide an important platform for multi-sectoral collaboration in integrated coastal and ocean management. Such collaboration facilitates knowledge creation and knowledge brokering at the science-policy interface. These processes can inform the

development of criteria for effective working relationships to promote open science (Eck, 2017; Soomai, Wells, & MacDonald, 2011).

d) Reports and documents prepared by environmental non-governmental organizations (eNGOS). Environmental non-governmental organizations often operate between governments and many stakeholder groups. These eNGOs "translate" and synthesize scientific reports and documents into accessible publications intended to inform and engage stakeholders and governments. The eNGOs function as boundary organizations and their publications serve as boundary objects. This activity and the related objects could serve as an example to open science programs that wish to promote transparency and produce information that is understandable by "all levels of an inquiring society." (Cadman, 2017; Cadman, MacDonald, & Soomai, in review).

e) Environmental assessments and Impact assessments
Environmental and impact assessments are tools used by governments at all levels (sub-national to international) for assessing potential environmental anthropogenic impacts. Such tools are a prominent example of where grey literature is relevant to open science. For decades, impact assessment processes have used grey literature in strategic, project-based deliberations, often involving a wide range of interested parties (Sadler &Dusík, 2016). In some jurisdictions (e.g., Canada), recent legislation has explicitly embraced open science as a policy and a methodology for mobilizing the grey literature generated within such assessment processes for collective democratic deliberation (Government of Canada, 2018). For relatively undeveloped contexts, e.g., Arctic regions, such new open science policies will have important consequences for how grey literature is used in open deliberative forums, including involving Indigenous communities (Stewart, 2018). The field of impact assessment and its publications may increasingly offer valuable test cases for considering how such grey literature contributes to open science mandates.

5. Open Access to Data

Large quantities of data are frequently produced in marine environmental science research conducted by governmental and intergovernmental organizations, such as from long-term monitoring programs that produce data over lengthy periods (Oppenheimer et al., 2019). Prior to advances in digital technologies, data files were often published in technical report series. Currently, data are maintained in digital files accessible through the internet. For instance, the Gulfwatch contaminants monitoring program of the GOMC has assembled large volumes of data and communicated findings to policy- and decision-makers in the Gulf of Maine and Bay of Fundy regions of the Northwest Atlantic since 1991 (Chase et al., 2001; personal observations of Wells). The Gulfwatch program has been a flagship initiative of the GOMC and the data have been released as raw data files, data reports, data summaries, and papers published in peer-reviewed journals (Chamberlain & Wells, 2014). The Gulfwatch data have provided considerable evidence for other reports and new research initiatives (Chamberlain, Wells, & MacDonald, 2018; Elskus et al., 2019).

To support use of the growing availability of large data sets, mapping tools such as digital coastal atlases have become important for information dissemination (O'Dea, Haddad, Dunne, & Walsh, 2011). Digital coastal atlases, as publicly available, web-based interactive tools, are increasingly valued by managers and other information users for their data visualization features. Participants in our study of marine atlases in Canada, Scotland, and the United States stated that coastal web-based atlases allowed them to be more confident about their decisions, since they were able to access and analyze a large volume of credible data in a centralized location (McLean, 2014). Because the data in the atlases are publicly available, the managers were able to explain their decision-making processes more easily to the public. Digital atlases offer another example of grey literature practices that model the expectations of open science.

6. Public Users of Grey Literature

Our research about information intended largely for environmental decision-making activities shows that the costs associated with the production of grey literature are usually borne by governmental and non-governmental organizations. The publications are freely available, making the information accessible to many levels of society and thereby encouraging engagement in public

policy decision processes. Assuming the content is understandable to a diversity of audiences, accessible grey literature promotes a community understanding of science. Science literacy is also stronger when availability is not restricted by paywalls or intellectual property constraints. Furthermore, accessible information products based on available scientific literature are a key means by which members of the interested public gather and disseminate marine knowledge, often later used in various aspects of the marine governance process (e.g., ocean literacy surveys, project consultations, legislative critique, community advisory committee participation, etc.). These observations are based on our recent study of place-based coastal values of members of a New England (USA) coastal community, pointing to the importance of understanding coastal perceptions (Ryder-Burbidge, 2017) and in a recent analysis of communication strategies employed by individual and non-governmental organization science communicators using two social media platforms (Martin, 2018; Martin & MacDonald (in review)). Free access, which is common with grey literature, is a feature emphasized in open science circles (Open science, n.d.) and the movement's push to make scientific information more accessible to all levels of societal inquiry is reflected both in science communication literature and by the interested public (Bickford et al., 2015; Martin, 2018; Martin & MacDonald, in review; Ryder-Burbidge, 2017; Soomai, MacDonald & Wells, 2013; Steel et al., 2005). Many ideas and innovations intended to open-up the scientific process, disseminate results therein, and effectively deploy analytic interpretations into policy-making processes can be drawn from recent research literature (Bickford et al., 2015; Lowndes et al., 2017; Martin, 2018; Martin & MacDonald, in review; Ryder-Burbidge, 2017; Wood-Charlson et al., 2015). Funding for these initiatives, however, continues to present an ongoing challenge. Our research suggests that financial models used in the production of grey literature could also inform open science models and help to sustain the continued development of knowledge translation projects.

Challenges about Grey Literature that Could Inform Open Science

Although produced and used widely in marine science (environmental and fisheries management) decision-making processes, grey literature presents challenges that open science also encounters (Table 2). Some challenges are long standing unresolved issues, whereas others are either mitigated or augmented by digital technologies.

Table 2. Challenges about Grey Literature that can Inform Open Science Practices

Challenges with Using Grey Literature Relevant to Open Science
Access (Physical and Understandability)
Awareness
Stigma about Data and Information
Production and Distribution Costs
Longevity of Websites

Examples of challenges include:

a) *Access*. While large quantities of grey literature are digitally available, various formats and platforms, etc., create access difficulties. A sizeable body of grey literature still exists in print formats only, or remains in proprietary digital formats, especially in the private sector. Moreover, when the grey literature is largely technical, the information may not be understandable by non-scientific audiences.

b) *Awareness*. The massive quantity of grey literature can result in reports or data hidden to potential users. Initiatives to promote awareness may be limited. In fact, there may be less incentive to promote awareness since "financial" rewards to promote awareness and use are not offered. One solution may be to encourage training in search strategies to find (or not find) the required information.

c) *Stigma*. The mistaken view that grey literature is always of lower quality than primary literature may also apply to open access (open science) primary literature. Today, this challenge is accentuated by the growing presence of predatory journals.

d) *Production and distribution costs borne by creators*. The production of grey literature has long followed the "creator pays" business model. While this model benefits users, i.e., the user does not pay to gain access to the literature, production and promotion costs, which are real, may pose a challenge for open science. Questions about who pays and who funds the production costs must be addressed.

e) *Longevity of the websites holding the data and the publications*. Website longevity is assumed but websites come and go, with information loss as the consequence. To date, no initiative has solved this problem. The Internet Archive, for example, captures large numbers of websites and other publications, but digital preservation of all websites has not been achieved (Baucom, 2019; Brügger & Laursen, 2019; Hill, 2016; Shein, 2016; Thelwall & Vaughan, 2004).

Openness does not ensure awareness and it does not automatically equate to usability by a wide variety of audiences. Because grey literature is assumed to be largely accessible, often limited attention is focused on promoting awareness. Furthermore, the massive quantity of literature can contribute to its seeming invisibility; specific publications may be unnoticed among the many available on a subject. The multiplicity of formats and varied content can result in grey literature being deemed of limited value. Even though the information may be rigorously peer-reviewed, in today's information-saturated environment, open-access may be equated with uncertain quality and questions of reliability (i.e., information vs. mis-information) and credibility arise.

Conclusions

Let us return to the question posed as a title of this paper, namely, "Is the production and use of grey marine literature a model for open science practices?" Our research suggests that it can be. We have observed and reported in publications arising from our studies that the long-standing motivations to produce grey literature and make it accessible are similar to the motivations driving open science. As seen in settings that explicitly invoke the use of grey literature in publicly deliberative moments (e.g., environmental impact assessments contexts), this similarity is increasingly being recognized (Government of Canada, 2018).

Proponents (creators) for both grey literature and open science aim to promote the use of information in order to advance research, to raise public awareness and understanding of science, and to facilitate evidence-based (evidence-informed) decisions on important societal issues.

The distinction between primary and grey literature is becoming less pronounced today, driven largely by advances in information technology and science generally. Often an immediacy to information needs overshadows the production processes. In public policy and decision processes, delays in making information available can be detrimental and costly.

Many definitions of open science are found in the literature on the subject. But a key observation from our research is that the distinction between grey and primary literature is blurring. Today, both types of literature emphasize greater open access. Moreover, it is important to note that significant marine environmental and fisheries research results must be published in a timely manner and distributed widely in formats suitable for diverse audiences.

Finally, we believe in the importance of ensuring that key marine information is published and made accessible, regardless of the location or the publishing method. Solving critical environmental problems by finding effective solutions and practicing sustainable fisheries should be the primary focus of research, management, and policy (Soomai & MacDonald, 2018). Attention given to the processes of information creation and delivery should be pursued with these urgent objectives in mind.

References

Allen, C., & Mehler, D. M. A. (2019). Open science challenges, benefits and tips in early career and beyond. *PLOS Biology, 17*(5), e3000246. https://doi.org/10.1371/journal.pbio.3000246

Baldwin, M. (2018). Scientific autonomy, public accountability, and the rise of "peer review" in the Cold War United States. *Isis, 109*(3), 538-558. https://doi.org/10.1086/700070

Baucom, E. (2019). A brief history of digital preservation. In J. Myntti & J. Zoom (Eds.), *Digital preservation in libraries: Preparing for a sustainable future* (pp. 3-19). Chicago: American Library Association.

Bickford, D., Posa, M. R. C., Qie, L., Campos-Arceiz, A., & Kudavidanage, E. P. (2012). Science communication for biodiversity conservation. *Biological Conservation, 151*(1), 74-76. https://doi.org/10.1016/j.biocon.2011.12.016

Blair, A. M. (2010). *Too much to know: Managing scholarly information before the modern age*. New Haven and London: Yale University Press.

Bohannon, J. (2013). Who's afraid of peer review? *Science, 342*(6154), 60–65 https://doi:10.1126/science.342.6154.60

Brügger, N., & Laursen, D. (Eds.). (2019). *The historical web and digital humanities: The case of national web domains.* London: Routledge.

Cadman, R. (2017). "Holding our feet to the fire": The role of ENGOs in environmental decision making. (Unpublished master's graduate research project report). Dalhousie University, Halifax, Canada.

Cadman, R., MacDonald, B. H., & Soomai, S. S. (in review). Sharing victories: Characteristics of collaborative strategies of environmental non-governmental organizations in Canadian marine conservation.

Cash, D., Clark, W. C., Alcock, F., Dickson, N., Eckley, N., & Jäger, J. (2002). Salience, credibility, legitimacy and boundaries: Linking research, assessment and decision making. *SSRN Electronic Journal*, [25]. https://doi.org/10.2139/ssrn.372280

Chamberlain, S. D., & Wells, P. G. (2014, October). Gulfwatch bibliography, Sept. 2014. Retrieved from www.gulfofmaine.org/public/wp-content/uploads/2017/06/GW-Outputs-Bibliography-D5-October-9-2014-Chamberlain-and-Wells.doc

Chamberlain, S. D., Wells, P. G., & MacDonald, B. H. (2018). The Gulfwatch contaminants monitoring program in the Gulf of Maine: Are its data used for ocean protection, with special reference to Nova Scotia, Canada? *Marine Pollution Bulletin, 127,* 781-787. http://dx.doi.org/10.1016/j.marpolbul.2017.09.050

Chase, M. E., Jones, S. H., Hennigar, P., Sowles, J., Harding, G. C. H., Freeman, K., … & Taylor, D. (2001). Gulfwatch: Monitoring spatial and temporal patterns of trace metal and organic contaminants in the Gulf of Maine (1991-1997) with the blue mussel, *Mytilus edulis L. Marine Pollution Bulletin, 42*(6), 491-505. https://doi.org/10.1016/S0025-326X(00)00193-4

Cordes, R. E. (2004). Is grey literature ever used? Using citation analysis to measure the impact of GESAMP, an international marine science advisory body. *Canadian Journal of Information and Library Science, 28*(1), 49-69.

Cossarini, D. M., MacDonald, B. H., & Wells, P. G. (2014). Communicating marine environmental information to decision makers: Enablers and barriers to use of publications (grey literature) of the Gulf of Maine Council on the Marine Environment. *Ocean & Coastal Management, 96,* 163-172. https://doi.org/10.1016/j.ocecoaman.2014.05.015

Eck, K. (2017). Evaluating the role of technical working groups in decision making for fisheries management in Belize. (Unpublished Master's graduate research report). Dalhousie University, Halifax, Canada.

Elskus, A. A., LeBlanc, L. A., Latimer, J. S., Page, D., Harding, G. C.H., & Wells, P. G. (in review). Monitoring chemical contaminants in the Gulf of Maine, using sediment and mussels: An evaluation. *Marine Pollution Bulletin.*

GESAMP (IMO/FAO/UNESCO/WMO/WHO/IAEA/UN/UNEP Joint Group of Experts on the Scientific Aspects of Marine Pollution). (1990). *The state of the marine environment*. Nairobi: United Nations Environment Programme. Reports and Studies, no. 39, and UNEP Regional Seas Reports and Studies, no. 115.

GESAMP (IMO/FAO/UNESCO-IOC/WMO/WHO/IAEA/UN/UNEP Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection). (2001). Protecting the oceans from land-based activities. Arendal, Norway: GRID-Arendal, United Nations Environmental Programme. Reports and Studies, no. 71

GESAMP (IMO/FAO/UNESCO-IOC/WMO/WHO/IAEA/UN/UNEP Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection). (2002). *Revised GESAMP hazard evaluation procedure for chemical substances carried by ships.* London: International Maritime Organization. Reports and Studies, no. 64.

Government of Canada. (2018, August 14). *The proposed impact assessment system. A technical guide*. Retrieved from https://www.canada.ca/content/dam/themes/environment/conservation/environmental-reviews/technical-guide.pdf

Haider, J., & Åström, F. (2017). Dimensions of trust in scholarly communication: Problematizing peer review in the aftermath of John Bohannon's "Sting" in science. *Journal of the Association for Information Science and Technology, 68*(2), 450-467. https://doi.org/10.1002/asi.23669

Hill, N. (2018). *Public record under threat: News and the archive in the age of digital distribution* (p. 13). Retrieved from Tow Center for Digital Journalism and the Brown Institute for Media Innovation website: https://academiccommons.columbia.edu/doi/10.7916/D8V99RMG

Huvila, I., Anderson, T. D., Jansen, E. H., McKenzie, P., & Worrall, A. (2017). Boundary objects in information science. *Journal of the Association for Information Science and Technology, 68*(8), 1807–1822. https://doi.org/10.1002/asi.23817

Landhuis, E. (2016, July 21). Information overload. *Nature, 535,* 457-458.

Lee, C. J., & Moher, D. (2017). Promote scientific integrity via journal peer review data. *Science, 357*(6348), 256–257. https://doi.org/10.1126/science.aan4141

Lowndes, J. S. S., Best, B. D., Scarborough, C., Afflerbach, J. C., Frazier, M. R., O'Hara, C. C., … Halpern, B. S. (2017). Our path to better science in less time using open data science tools. *Nature Ecology & Evolution, 1*(6), 0160. https://doi.org/10.1038/s41559-017-0160

MacDonald, B. H., Cordes, R. E., & Wells, P. G. (2004). Grey literature in the life of GESAMP, an international marine scientific advisory body. *Publishing Research Quarterly, 20*(1), 25-41.

MacDonald, B. H., Cordes, R. E., & Wells, P. G. (2007). Assessing the diffusion and impact of grey literature published by international intergovernmental groups: The case of the Gulf of Maine Council on the Marine Environment. *Publishing Research Quarterly, 23*(1), 30-46. doi:10.1007/s12109-007-9010-6

MacDonald, B. H., De Santo, E. M., Quigley, K., Soomai, S. S., & Wells, P. G. (2013). Tracking the influence of grey literature in public policy contexts: The necessity and benefit of interdisciplinary research. *The Grey Journal, 9*(2), 61-69.

MacDonald, B. H., Soomai, S. S., De Santo, E. M., & Wells, P. G. (Eds.). (2016). *Science, information, and policy interface for effective coastal and ocean management*. Boca Raton: CRC Press (Taylor & Francis).

MacDonald, B. H., & Soomai, S. S. (2019). Environmental research and knowledge production within governmental organizations. In L. Bőrjesson & I. Huvila (Eds.). *Research outside the academy: Professional knowledge-making in the digital age* (pp. 21-50). Cham: Palgrave Macmillan.

Martin, C. (2018). "It feels like engaging with a friend": Using interpersonal communication strategies to encourage science conversations with lay audiences on social media. (Unpublished master's graduate project report). Dalhousie University, Halifax, Canada.

Martin, C., & MacDonald, B. H. (in review). Using interpersonal communication strategies to encourage science conversations in social media.

McLean, S. (2014). *A study of the use of data provided by coastal atlases in coastal policy and decision-making.* (Unpublished Master's project). Dalhousie University, Halifax, Canada.

Méndez, E. (2019). Open science? …Darling, we need to talk. Open Science Conference, 19-20 March 2019, Berlin. Retrieved from https://www.open-science-conference.eu/wp-content/uploads/2019/03/Eva-Mendez.pdf

Muffler-Wille, S., & Charmantier, I. (2012). Natural history and information overload: The case of Linnaeus. *Studies in History and Philosophy of Biological and Biomedical Sciences, 43*(1), 3-15.

O'Dea, E., Haddad, T. C., Dunne, D., & Walsh, K. (2011). Coastal web atlas features. In D. Wright, N. Dwyer, & V. Cummins (Eds.). *Coastal informatics: Web atlas design and implementation* (pp. 12-32). Hershey, PA: IGI Global. doi: 10.4018/978-1-61520-815-9.ch002

Ogilvie, B. W. (2003). The many books of nature: Renaissance naturalists and information overload. *Journal of the History of Ideas, 64*(1), 29-40.

Open science. (n.d.). Retrieved from http://www.oecd.org/science/inno/open-science.htm

Open science. (2019, September 19). In Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Open_science

Open science definition. (n.d.). Retrieved from https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition

Oppenheimer, M., Oreskes, N., Jamieson, D., Brysee, K., O'Reilly, J., & Shindell, M. (2019). *Discerning experts: The practices of scientific assessment for environmental policy.* Chicago: University of Chicago Press.

Rosenberg, D. (2003). Early modern information overload. *Journal of the History of Ideas, 64*(1), 1-9.

Ross, J. D. (2014). What do users want from a state of the environment report? A study of the awareness and use of the *State of the Scotia Shelf Report.* (Unpublished Master's thesis). Dalhousie University, Halifax, Canada.

Ross, J. D., & Breeze, H. (2016). What do users want from a state of the environment report? A case study of awareness and use of Canada's *State of the Scotian Shelf Report*. In B. H. MacDonald. S. S. Soomai, E. M. De Santo, & P. G. Wells (Eds.). *Science, information, and policy interface for effective coastal and ocean management* (pp. 283-302). Boca Raton: CRC Press (Taylor & Francis).

Ryder-Burbidge, S. (2017). "I thought the horseshoe crabs were part of my family": Investigating ocean connectivity in Falmouth, Massachusetts. (Unpublished master's graduate project report). Dalhousie University, Halifax, Canada.

Sadler, B., & Dusik, J. (Eds.). (2016). *European and international experiences of strategic environmental assessment*. London: Earthscan.

Shein, E. (2016). Preserving the internet. *Communications of the ACM, 59*(1), 26-28. https://doi.org/10.1145/2843553

Soomai, S. S. (2017a). The science-policy interface in fisheries management: Insights about the influence of organizational structure and culture on information pathways. *Marine Policy, 81*, 53-63. https://doi.org/10.1016/j.marpol.2017.03.016

Soomai, S. S. (2017b). Understanding the science-policy interface: Case studies on the role of information in fisheries management. *Environmental Science & Policy, 72*, 65-75. https://doi.org/10.1016/j.envsci.2017.03.004

Soomai, S. S., & MacDonald, B. H. (2018). Information matters: Global perspectives about communication at the science-policy interface. In D. Werle, P. R. Boudreau, M. R. Brooks, M. J. A. Butler, A. Charles, S. Coffen-Smout, … P. G. Wells (Eds.), *The future of ocean governance and capacity development: Essays in honor of Elisabeth Mann Borgese (1918-2002)* (pp. 271-276). https://doi.org/10.1163/9789004380271

Soomai, S. S., MacDonald, B. H., & Wells, P. G. (2013). Communicating environmental information to the stakeholders in coastal and marine policy-making: Case studies from Nova Scotia and the Gulf of Maine/Bay of Fundy region. *Marine Policy, 40,* 176-186.

Soomai, S.S., Wells, P. G., & MacDonald, B. H. (2011). Multistakeholder perspectives on the use and influence of "grey" scientific information in fisheries management. *Marine Policy, 33*(1), 50-62. doi:10.1016/j.marpol.2010.07.006

Steel, B., Smith, C., Opsommer, L., Curiel, S., & Warner-Steel, R. (2005). Public ocean literacy in the United States. *Ocean and Coastal Management, 48*(2), 97-114. https://doi.org/10.1016/j.ocecoaman.2005.01.002

Stewart, I. (2018). From environmental assessment to impact assessment under Bill C-69: Some science policy implications of changing landscape of federal impact assessment for offshore O&G. Presented at WWF Arctic Oil and Gas Symposium, Ottawa, September 2018.

Thelwall, M., & Vaughan, L. (2004). A fair history of the Web? Examining country balance in the Internet Archive. *Library & Information Science Research, 26*(2), 162-176. https://doi.org/10.1016/j.lisr.2003.12.009

Tidline, T. J. (2002). Information overload. In A. Kent, & C. M. Hall (Eds.). *Encyclopedia of Library and Information Science* (72, supplement 35, pp. 217-234). New York: Marcel Dekker, Inc.

Vermeir, K., & Margócsy, D. (Eds.). (2012). States of secrecy. Special Issue. *The British Journal for the History of Science, 45*(2), 153-280. doi:10.1017/S0007087412000052

Vincente-Saez, R., & Martinez-Fuentes, C. (2018). Open science now: A systematic literature review for an integrated definition. *Journal of Business Research, 88*, 428-436. https://doi.org/10.1016/j.jbusres.2017.12.043

Walgrave, S., & Dejaeghere, Y. (2017). Surviving information overload: How elite politicians select information. *Governance, 30*(2), 229-244. https://doi.org/10.1111/gove.12209

Wells, P. G. (2003). State of the marine environment reports -- a need to evaluate their role in marine environmental protection and conservation. *Marine Pollution Bulletin, 46*(10), 1219-1223. doi:1 0.1016/S0025-326X(03)00284-4

Wells, P. G. (2016). A career-based perspective of science-policy linkages in Environment Canada: The role of information in managing human activities in our ocean spaces. In B. H. MacDonald, S. S. Soomai, E. M. De Santo, & P. G. Wells (Eds.). *Science, information, and policy interface for effective coastal and ocean management* (pp. 367-388). Boca Raton: CRC Press.

Wells, P. G., Duce, R. A., & Huber, M. E. (2002). Caring for the sea — accomplishments, activities and future of the United Nations GESAMP (the Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection). *Ocean & Coastal Management, 45,* 78-89. https://doi.org/10.1016/S0964-5691(02)00047-9

Woelfle, M., Olliaro, P., & Todd, M. H. (2011). Open science is a research accelerator. *Nature Chemistry, 3*(10), 745-748. doi:10.1038/nchem.1149.

Wood-Charlson, E. M., Bender, S. J., Bruno, B. C., Diaz, J. M., Gradoville, M. R., Loury, E., & Viviani, D. A. (2015). Translating science into stories. *Limnology and Oceanography Bulletin, 24*(3), 73-76. https://doi.org/10.1002/lob.10055

Abstracting and Indexing as an enabling interface between open science and grey literature – The approach of the Aquatic Sciences and Fisheries Abstracts service*

Tamsin Vicary, Food and Agriculture Organization of the United Nations, Italy
Ian Pettman, Freshwater Biological Association, United Kingdom

Abstract:
We examine the role of one Abstracting and Indexing (A&I) service (Aquatic Sciences and Fisheries Abstracts - ASFA) as an interface between open science and grey literature. As an A&I service known for its coverage of grey literature, ASFA is evolving its technologies and partnerships to capitalize on the opportunities presented by Open Science. A Grey Literature strategy was implemented to ensure ASFA made the necessary changes to its monitoring, recording and reporting of grey literature coverage on its database. We describe how the strategy incorporates the Open Science movement into ASFA, by making use of the opportunities such as the increased number of repositories and literature; and vice-versa by providing increased access to grey literature and working with institutions to ensure their literature is captured on ASFA. Specifically we describe how ASFA is: (1) increasing access to grey literature by working in partnership with OA repositories (including harvesting from OAI-PMH compliant repositories); (2) using open source software to manage its indexing and search tools, allowing them to be utilized by, and receive contributions from, a greater number of users to aid the discoverability of grey literature; (3) ensuring comprehensive and timely coverage of grey literature among its global partnership, and providing funding for small specific projects which meet ASFA's grey literature priorities regards subject scope and content type; (4) working with the Open Science community in a Virtual Research Environment (VRE) – such as iMarine – to ensure a two way interaction between the Open Science movement and the promotion of grey literature in the future. We demonstrate that by taking the above steps to capitalize on the opportunities presented by Open Science and new technologies, an A&I service such as ASFA becomes a valued interface between Open Science and grey literature – ensuring the recording and discovery of grey literature from aquatic sciences and fisheries institutions around the world.

Introduction

The ASFA database was first published as a monthly abstract journal in July 1971. An international cooperative, input to the database is today provided by 64 national and international partners, 43 collaborative centres, and the commercial publisher ProQuest. Each of ASFA's Partners has signed an agreement with FAO, stating that they will be responsible for monitoring serials and non-conventional (grey) literature relevant to the scope of ASFA published in their own countries and for preparing bibliographic citations, indexing and abstracts of relevant literature for input to ASFA. In the pre-digital age of ASFA's origins, this model ensured both primary and non-conventional and grey literature was adequately covered. For example, in 1979, ASFA Board members estimated that 40,000 publications within ASFA's subject scope were issued annually, and set ASFA a medium-term target of covering 30,000 records, so roughly 75% of all relevant literature. As the commercial publisher concentrated on primary literature, GL input was provided by ASFA Partners who in 1974 provided 50% of records added to the database[1]. In recent years, that percentage has changed in favour of ProQuest who, in the last five years has provided an average of 86% of the input to the ASFA database (57,6693 records out of 63,7428 records in total, from 2014-2018). As ProQuest continues to concentrate on monitoring primary literature, ASFA's GL coverage remains dependent on input from its Partners. Without standard reporting and analytics, the actual composition of GL being added to the database was unknown. Whilst ASFA's historic coverage of GL remains undisputed and a key way of differentiating it from other information products, has ASFA kept pace with the digital age and influx of material resulting from Open Science initiatives to still adequately cover GL as it did in its early years? Has the ASFA Partnership adapted its technologies to react to

* First published in the GL21 Conference Proceedings, February 2020.
[1] Varley (1995)

and benefit from opportunities provided by Open Science, and how should its business model change to reflect changing user expectations? These are questions we attempted to address, whose answers resulted in a strategy to improve the coverage of GL on ASFA as well as recommending revisions to ASFA's business model to better meet Open Science principles[2].

Foremost among these revisions, is increasing access to ASFA information products and services. In 2018, FAO implemented an Open Access policy, advocating "the application of suitable open licences to FAO copyright material in accordance with the principles of openness and sharing envisioned under Open Access, and consistent with the mandate of FAO.[3]" All ASFA Products (including the database) are owned by FAO, and as a project supported by FAO, ASFA must comply with FAO policies and move to ensure openness of its products. ASFA must also adapt to increase its support to FAO strategic objectives and the UN Sustainable Development Goals.[4] Monitoring progress on all SDGs "requires constant scientific input and would not be possible without opening access to relevant data"[5] – meaning that by enabling access to relevant data ASFA has a role to play in supporting SDG 14, Life Below Water.

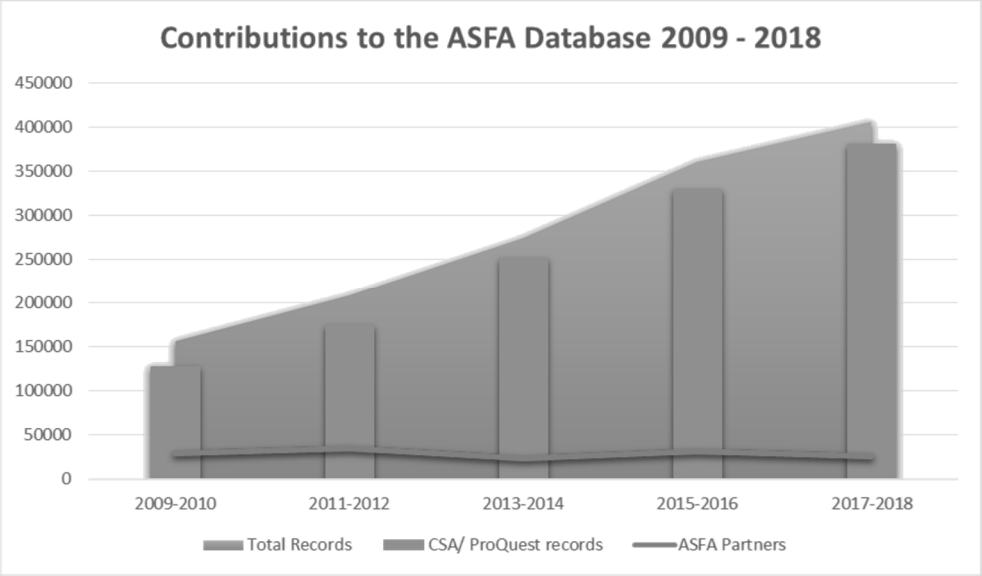
Open Science therefore presents both challenge and opportunity to ASFA, and has forced ASFA to adapt and improve its policies and technologies, both learning from and becoming an educator on the benefits of GL to Open Science to both authors and users. Together, these changes have led to the formation of a GL strategy the embraces Open Science principles and delivers the benefits to ASFA Partners and database users. However before addressing the changes needed, it is first necessary to assess ASFA's GL coverage to understand its strengths and weaknesses and in order to provide a baseline to which future improvements can be measured against.

Assessment of ASFA's GL Coverage

As noted, ASFA has a reputation as a provider of GL, however little has been done to define or quantify coverage on ASFA, or to outline how ASFA Partners should monitor GL. At the 2018 ASFA Advisory Board Meeting (Ostend, Belgium), Peter Pissiersiens (Head, IOC Project Office for IODE) commented on ASFA's GL coverage, asking whether the increase in available literature had been matched by an increase in records being added to the ASFA database, and whether ASFA had a definition for grey literature. The ASFA Secretariat took the first tentative steps in making this assessment.

The below graph shows that the total number of records added to the ASFA database has steadily increased since 2009, however that the increase is due to ProQuest adding more records to the database and not due to an increase in ASFA Partner efforts, which have remained stable. Although the number has fluctuated there has been no significant change in the number of ASFA Partners, therefore the number of records per partner has stayed roughly the same from 2009 -2018. From this we can conclude that any increase in available information in the last ten years has not been matched in an increase by Partners, however, this does not mean that Partners have not increased their GL input, as it may be that ProQuest has taken on monitoring responsibility for commercially published journals in this period, allowing Partners to focus more on grey literature. We therefore need to assess Partners' grey literature input to the database.

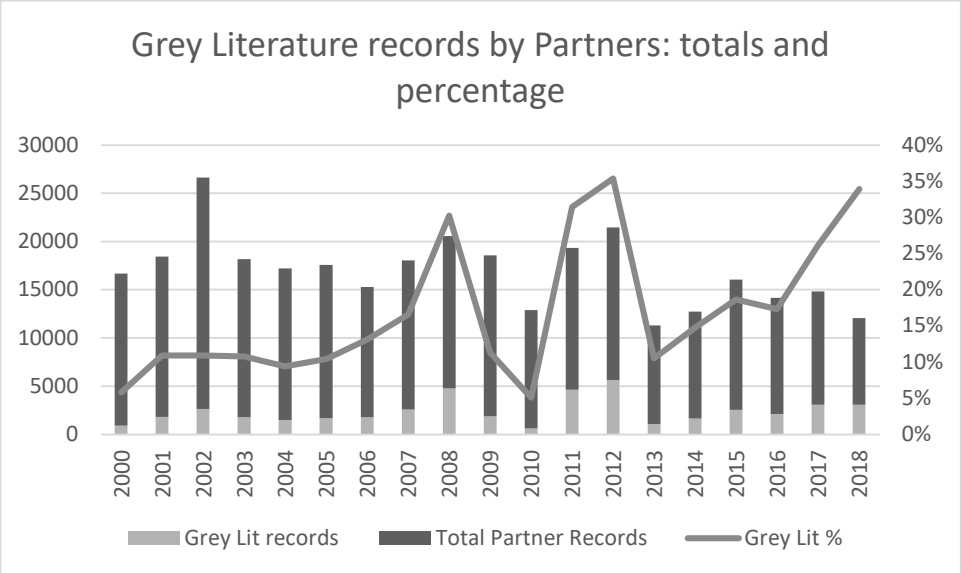
[2] Open and Collaborative Science in Development Network (2017)
[3] FAO (2018)
[4] United Nations (2019)
[5] United Nations (2018)



In order to get an indication of how much grey literature is being added to the database, we used the below search string, based on the Source type metadata field. Although not fully accurate, by searching defined source types we can at least measure how much literature that falls under the GL framework is being added to the database by Partners.

Search string (for 2016):
Input centre: All ASFA Partners, excluding ProQuest records
Source type: Conference Papers and Proceedings; Dissertations and Theses; Government & Official Publications; Reports; Working Papers; Other sources.

As the below graph shows, the percentage of GL added to the database as measured by Source Type has fluctuated, from a low of 621 records (5% of records) in 2010, to a high of 5605 records (35% of Partner records) in 2012.



According to these figures, there has been an increase in both the percentage and total number of GL records added by Partners since 2013. ASFA Partner GL records rose from 1073 grey literature records (10%) in 2013, to 3055 grey literature records (34%) in 2018. Though the period 2008 - 2013 fluctuated, the increase since 2013, and its comparison to 2000 - 2007, indicates that ASFA Partners coverage of GL has increased, when measured by percentage of total records. ProQuest has taken on responsibility to cover primary journals that were previously monitored by Partners, which enables Partners to spend more time monitoring GL. The potential of ProQuest further

increasing its primary journal coverage to allow Partners to concentrate on GL was discussed and will form part of ASFA's strategy to improve its GL coverage.

To assist Partners with monitoring the GL in their countries, a discussion session was held at the ASFA Advisory Board Meeting (Malaysia, 22-25 September 2019), to define what constituted GL for Partners, and what would be most useful to cover for ASFA. A quiz was held, which asked Partners what they believed constituted GL and what did not, with the following results showing significant divisions among partners:

	Technical Guidelines	**Video – online course**	**Infographic**	**Dissertations / Theses**	**Journal article in *Nature***	**Monograph by university press**
Yes	11	13	21	18	4	3
No	6	11	2	1	19	22

Lack of understanding of what constitutes GL by Partners is likely to lead to them not monitoring GL for ASFA, and not being able to report on their coverage. ASFA has initiated several knowledge sharing initiatives among ASFA Partners and beyond, in order to improve understanding. These include the ASFA Newsletter (Issue 2 focused on Grey Literature) and a Conference on GL which was held in Malaysia after the Board Meeting (discussed below). It is hoped these measures will ensure Partners better understand the importance of GL and will look to include more on ASFA.

Following the quiz, Partners discussed and agreed a working definition of GL and decided to prioritize certain content types to help improve coverage of GL on the ASFA database. The following definition, based on the Prague definition, was designed to be broad enough to include a variety of GL content types produced by various GL publishers, and specific enough to be of use to ASFA Partners and database users:

> *Grey literature is information presented in any number of physical or digital formats, under the subject scope of aquatic sciences, fisheries or aquaculture, of sufficient quality to be preserved and of public good but produced outside the control of commercial publishers.*

Reports by Partners and discussions at ASFA's 2019 Advisory Board meeting concluded that the following content types would be prioritized:

> Conference Papers, Technical Reports, Research Reports and Dissertations and Theses.

Furthermore, two ASFA Partners (the FBA in UK and National Fisheries Resources Research Institute in Uganda) indicated that their institutions have data sets which are currently seeking a permanent storage solution. Following discussions, the decision was made to explore creating an OpenAIRE community that would store these data sets and be used as a pilot study to assess whether ASFA can incorporate grey data as well as literature. Much historical data on aquatic science is included in grey literature reports, often in non-digital format and residing in libraries - making this data accessible is essential to "deciphering past and current trends in environmental conditions and populations of living re-sources" (Wells, 2014).

Providing knowledge sharing activities, a working definition, prioritizing content types and exploring ways to cover new content types such as data, will provide a framework to ASFA Partners who are tasked with covering the GL of their institute, helping them to identify and record GL of value to end users. However, this framework will only be of us if ASFA technologies are updated to provide Partners with the necessary tools to efficiently capture and disseminate the literature in their country or region.

Technology as a driving force for ASFA's adoption of Open Science principles

According to OpenAIRE, Open Science has created a 'new modus operandi for science where stakeholders…are involved and research is organized, linked, verified, facilitated by new technologies and enhanced with collaborative and coordinative activities.'[6] The development and adoption of new technologies has undoubtedly enabled the Open Science movement to increase the volume and content type made available throughout the research cycle, clearly benefitting GL due to technologies such as repositories, search engines, Linked Open Data and controlled vocabularies. Together, these technologies have provided platforms and discovery services for many GL types that would previously have remained hidden on institutional networks or premises, leading to an explosion of available literature and data – in the last five years, the number of repositories on OpenDOAR has grown from 2743 to 4358[7] and the number of Science and Engineering articles has grown by an average of 3.9% each year between 2006 and 2016, according to the National Science Foundation, USA[8]. Whilst aggregation services have been successful in enabling searching of these growing number of heterogeneous repositories through one interface, the growth in the number of OA documents is now beginning to make retrieval precision a challenge for the open science movement – the aggregator CORE has 135,539,113 open access records (https://core.ac.uk/) but its simple metadata does not allow for advanced searching. A subject specific discovery service such as ASFA, which is able to provide precise retrieval through its detailed metadata structure and search platform, could have a role to play in Open Science by incorporating these new technologies in order to capture the increasing volumes and various document types being produced, placing them on a domain specific platform and enhancing retrieval by improving the metadata quality.

At the 2019 ASFA Advisory Board Meeting it was decided to investigate the potential of collaboration with the OpenAIRE aggregator and deposit service in order to improve the coverage and precision of recall for GL in the aquatic subject domain. An example may be the possibilities of an "Aquatic Community" area within OpenAIRE – which ASFA could potentially advance in collaboration with Aquatic Commons, OceanDocs and other relevant repositories in order to improve exposure of the GL. It is worth noting here, that not all technologies have been adopted by institutions equally across the world; many institutions in developing countries still lack repositories and infrastructure to adequately publish, store and disseminate their information. Of the top ten countries by number of repositories on OpenDOAR, only two (Peru and Brazil) are classified as developing countries, and together these represent less than 12% of the total number of repositories covered by the top ten countries[9]. In the past, ASFA has funded digitization projects in developing countries, providing both hardware and access to digital infrastructure. ASFA will continue to provide these services, however it will also look to increase its impact, particularly in developing countries, by adopting two technologies: (1) a Virtual Research Environment to handle the creation (by both manual input and harvesting), storage and publishing of records; and (2) open source software VocBench which is used to manage and maintain the ASFA Subject vocabulary. We argue that by adopting these technologies, ASFA becomes an interface between the growing volume of literature being produced under Open Science principles and the information users seeking specific, relevant and credible information.

Virtual Research Environment

Having used DOS based CDS ISIS for creating records in .ISO format, ASFA is now moving to an online system, a Virtual Research Environment (VRE). The VRE will provide ASFA partners with a single space to manage the creation and publishing of ASFA's bibliographic records, the import and harvesting of metadata from other repositories, the export of records to the ASFA publisher ProQuest, and the provision of API services to websites and services, including the FAO Fisheries and Aquaculture website, in order to integrate ASFA records with other platforms. Formed by a

[6] OpenAIRE, (2019)
[7] OpenDOAR (2019)
[8] National Science Foundation (2018)
[9] OpenDOAR (2019)

need to adapt to the changes Open Science engendered, the VRE services will contribute themselves to Open Science. The records Partners create will now be openly searchable and accessible on the VRE, the first time ASFA has provided an open platform outside of the commercial database published by ProQuest. Through API services, records can display on an institutional website with trials showing how the subject, geographic and taxonomic keywords can be used to link ASFA to existing FAO Fisheries and Aquaculture information systems, such as country profiles and stocks data at species level.

To facilitate its move to the VRE, ASFA mapped its metadata fields to Dublin Core, ensuring a high degree of interoperability with repositories such as OceanDocs. ASFA will harvest metadata from 10 ASFA Partners' repositories, which are OAI-PMH compliant. If successful this will be expanded to a greater number of repositories. The ASFA Software Working Group is presently reviewing our mappings to ensure, where possible, compatibility with repository metadata standards such as RIOXX and the OpenAIRE metadata fields. Due to ASFA's detailed metadata, harvested feeds will need to be monitored and edited by the responsible Partner, for example the insertion of taxonomic and geographic keywords. Though Open Science has increased the volume of available literature, metadata in many institutional repositories remains simple therefore limiting its recall and interoperability. Investigations are underway to automate this procedure, however as an A&I service, ASFA's high quality and detailed metadata is key to differentiating itself from other information products, therefore although automation is useful for quickening the process, a degree of human editing will always be important for ASFA to maintain its values of accurate and detailed metadata. This is particularly true on many university repositories, where a large degree of manual filtering is required in order to retrieve relevant results. Over the last two years, ASFA has been exploring some of the challenges in locating open access full text university theses relevant to the aquatic community. Some of the preliminary findings have been outlined in ASFA Newsletter issue 2, 2019[10].

ASFA's metadata structure is an investment and should not be limited to closed access, subscription only products. Therefore through the use of its VRE, ASFA is exploring ways to utilize its metadata to contribute and collarborate with other products and services. For example, plans are underway to use ASFA's taxonomic keywords to link to specific species and strains of Aquatic Genetic resources in a registry. An open, interoperable platform was initially seen as essential to increasing ASFA's individual presence and prestige, however having made the decision to move to an open platform, it is proving to be an assured avenue to collaborating and enhancing FAO information products and services, thereby ensuring compliance with FAO's Open Access policy.

ASFA Subject Vocabulary

ASFA is currently using the open source software VocBench to manage and maintain its subject vocabulary, used for indexing ASFA records. The vocabulary is an RDF/SKOS-XL concept scheme, and a Linked Open Data (LOD) set. It is available online on SKOSMOS to search and browse; as a file to download; via web services. ASFA has a working group of 15 members to manage and maintain the vocabulary, each of whom is provided with training and support. The VRE links to the vocabulary via web services, and the ASFA Secretariat is also working with AGROVOC (FAO's main vocabulary which covers all agricultural elements), aligning concepts to improve the fisheries and aquaculture coverage.

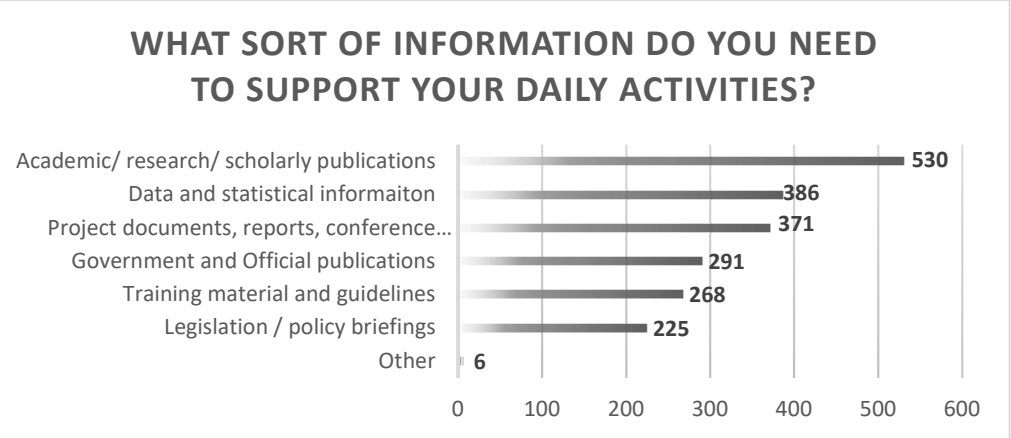
A well maintained thesaurus provides users with the option to filter searches and retrieve results with a higher level of specificity, which is a key way for ASFA to enhance the retrieval of the increasing volumes of GL that Open Science has enabled. The multi-lingual capability of the vocabulary will benefit both inputters creating ASFA records, and non-English speaking users when used in non-English repositories. The ASFA Subject Vocabulary thereby meets the Open Science principle of incentivizing "inclusive infrastructures that empower people of all abilities to make, and use accessible open-source technologies" (Open and Collaborative Science in Development Network, 2017).

[10] Pettman, I. (2019)

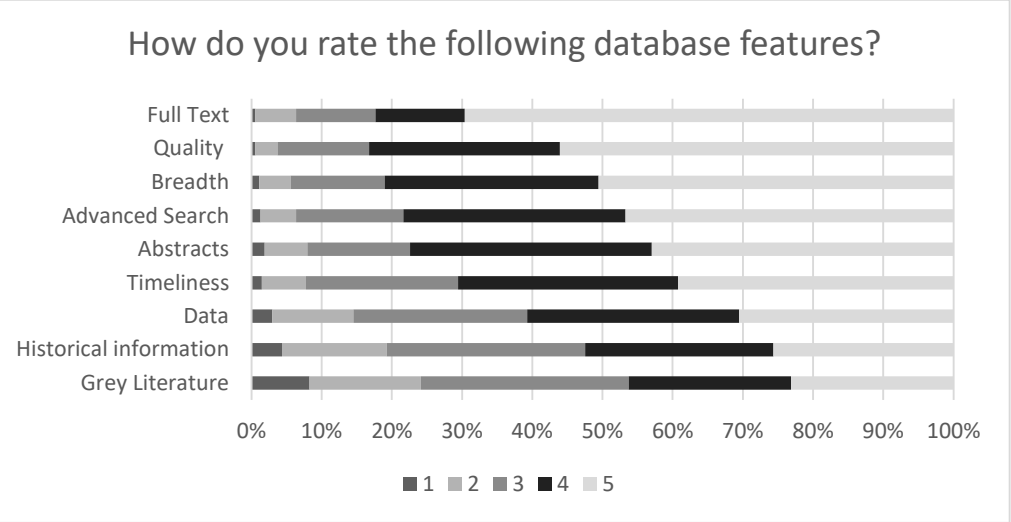
Through its VRE and Subject Thesaurus, ASFA is enabling the interaction between Open Science and GL as both of these technologies enable precise recall of aquatic sciences, fisheries and aquaculture information – but technologies alone do not ensure ASFA or Open Science meet their goals of increasing accessibility of literature. Education, the promotion of the value of grey literature to users and authors, is also necessary to ensure these technologies are used, and therefore must also be considered by ASFA as it formulates its Grey Literature strategy.

ASFA as an educator of Open Science benefits to GL users and publishers

During the course of a number of impact evaluation exercises, ASFA has attempted to assess user needs in the fields of aquatic sciences, fisheries and aquaculture. Having consulted those both inside and outside of the ASFA Partnership, it became clear that the term 'Grey Literature' was poorly understood and defined by both groups. For instance, GL content types were the third most required type of information for user survey respondents, however the inclusion of GL was least likely to be rated as an important database feature by the same respondents - this option also received the smallest number of 9 options presented to responders (total of 497 ranked inclusion of GL versus an average of 535 who ranked the other 8 options). This indicates there could be a lack of understanding on what actually constitutes GL.



Versus:



Lack of understanding by authors of GL is likely to lead them to not publish or make available their own GL. Educating both ASFA Partners and GL authors on the importance of GL is likely to help avoid unnecessary loss of information. ASFA has a role to play in educating users and authors of GL on the value of their information and encouraging deposition on an openly searchable repository, with the final step of adding a bibliographic record to the ASFA database to enable discover on a global platform.

Promoting the benefits of GL storage, publishing and dissemination

Information presented at the joint ASFA-UMT conference on GL in aquatic sciences (held 25th September 2019 at University Malaysia Terengganu) provided a number of ways for ASFA to promote the benefits of making their GL accessible to a wide audience. Firstly, a presentation by Dr Amirrudin B. Ahmad, of UMT, Malaysia entitled 'Mainstreaming Grey Literature in the Digital Age'[11] highlighted the importance of ensuring publication standards for GL whilst avoiding lengthy and unnecessary delays due to publication control. Dr Amirrudin stressed that the benefits of making GL available on an international platform such as ASFA, especially to students who are currently progressing through their dissertations and phd theses, saying that ASFA provided access to relevant, domain specific information in a precise way not achievable by using search engines. ASFA will continue to work with Dr Amirrudin to produce a 'Top Tips' guide to publishing grey literature, which will be openly available and share knowledge and expertise not only on why but how to publish GL.

A second presentation presented a bibliometric analysis of GL usage among Filipino aquaculture researchers. Based on citation analysis, it was found that:

- There is a higher number of GL citations when first author is Filipino and / or when journal is published in the Philippines
- Only 25% of top GL publishers included on ASFA
- Websites emerged as one of most cited sources

The presentation concluded by recommending that non-traditional sources were included on ASFA and that ASFA Partners seek to increase the number of publishers they monitor for AFSA. The analysis proved a useful way to identify GL of use to researchers and ASFA, and the methodology will be replicated in other ASFA Partner countries to form a global assessment of GL publishers and ensure better coverage on the database. This work will be coordinated by ASFA's Impact and Strategies Working Group.

Supporting Open Science by promoting Grey Literature

FAO's mandate is to "collect, analyse, interpret and disseminate information relating to nutrition, food and agriculture". Having provided the ASFA Secretariat since 1975, it is incumbent on ASFA to demonstrate how it meets supports FAO's mandate. Clearly, ASFA's support to a subscription-only database is not sufficient support to FAO's goals, and ASFA is in the process of revising its Business Model to revise its functions and priorities. This section covers the reasons why ASFA has chosen to focus its strategy on improving GL coverage, and how by doing so ASFA is meeting not only FAO goals, and how Open Science has influenced ASFA's strategic direction. Having developed the technologies to enable an open platform, ASFA will be contributing to two FAO projects to increase access to information.

1. PESCAO

ASFA is participating in FAO project PESCAO to make an inventory of primary and GL in CECAF (Fishery Committee for the Eastern Central Atlantic) countries. The PESCAO-CECAF Project is to improve regional governance of marine resources in the CECAF area using knowledge-based advice by strengthening the management processes to contribute to sustainable fisheries, food security, and sustainable livelihoods. The inventory will be openly accessible and contribute to identifying research strengths and weaknesses across the CECAF region, thus contributing to the development of future fisheries research projects.

ASFA will use the exercise to identify and fill gaps, storing records on OA repositories. Use of Partners expertise and networks key to assessing research, as well as technologies to capture and store GL.

[11] Dr Amirrudin B. Ahmad (2019)

2. Aquatic Genetic Resources (AqGR) Registry

ASFA is ideally placed to contribute to growing knowledge and awareness of aquaculture resources. ASFA's subject thesaurus helps identify information of interest to AqGR, and the geographic and taxonomic terms can be used to specify where aquaculture research is taking place worldwide and on what species. For example, when an ASFA Partner creates a record on the genetics of Clarias gariepinus used for aquaculture in Nigeria, this can be linked to the appropriate record in the AqGR registry. Combining the ASFA bibliographic records with data in the AqGR registry therefore ensures a highly specific information stream on aquaculture species, alerting users of research conducted by research institutions, NGOs, academia, who all contribute to ASFA.

Though much scientific literature is openly available online, ASFA's use of controlled vocabulary terms to index its records mean it can deliver a level of accuracy and specificity to data and information systems such as AqGR.

Conclusion

By increasing access to information, Open Science has increased the expectations of users and forced ASFA to adapt its products and services. By adapting its technologies to provide openly accessible information products, such as its VRE and Subject Thesaurus, ASFA is demonstrating the value of an A&I service in the digital age; for example ASFA's detailed metadata can be used to enhance other information products and also lead to better search precision, an aspect which is likely to be of increasing importance as more and more information becomes available. Detailed metadata should not be the preserve of commercial only products and by collaborating and enhancing with other information products, as well as providing an open platform itself, ASFA is going some way to ensure the benefits of detailed metadata and its investment in new technologies, reach their maximum potential use.

Though A&I services are seen by many as old fashioned or obsolete, we believe that by embracing new technologies and policies as advocated by Open Science and Open Access policies, ASFA can ensure relevancy and usefulness in the digital age. ASFA's international partnership model ensures these technologies and information products are available to institutions worldwide. By providing a reputable and global database, ASFA helps to counter publication bias and ensures that valuable research performed in countries and regions underrepresented by traditional publishers is not lost to users.

In addition to its metadata and technologies, ASFA also has a role to play in knowledge sharing and promoting the benefits of GL. It is hoped that by interacting with its various stakeholders (ASFA Partners, database users, GL authors), ASFA will increase the amount of literature being captured and used in the fields of aquatic sciences, fisheries and aquaculture. This will ensure that the valuable information and data held in many GL reports around the world contributes to understanding and solving the critical problems facing the world's oceans and freshwater environments.

References

Ahmad, A.B. (2019) *Mainstreaming Grey Literature in the Digital Age.* Paper presented at Joint ASFA-UMT Conference Enabling Grey Literature Discovery Benefits Aquatic Sciences, Fisheries and Aquaculture Research, University of Malaysia Terengganu.

Candela, L., Castelli D., Pagano, P. (2013) Virtual Research Environments: An Overview and a Research Agenda. *Data Science Journal*, 12, pp. GRDI75-GRDI81 DOI: http://dx.doi.org/10.2481/dsj.GRDI-013

FAO (2017) Basic texts of the Food and Agriculture Organization of the United Nations. Volumes I and II. Rome, Italy: FAO.

FAO (2018). *FAO Open Access Policy.* Rome, Italy: FAO.

Messaoudi, S., Endra, A., Okeyo, R. (2019) *Valorization of Grey Literature in Aquatic and Marine Libraries in Africa by its Submission and Discovery via new technologies: Aquatic Commons, OceanDocs and ASFA.* Paper presented at Joint ASFA-UMT Conference Enabling Grey Literature Discovery Benefits Aquatic Sciences, Fisheries and Aquaculture Research, University of Malaysia Terengganu.

National Science Board (2018) *Science and Engineering Indicators 2018*. NSB-2018-1. Alexandria, VA: National Science Foundation. Available at https://www.nsf.gov/statistics/indicators/

Open and Collaborative Science in Development Network (2017) *Open Science Manifesto.* Retrieved from https://ocsdnet.org/manifesto/open-science-manifesto/

OpenAIRE. (2019). *Open Science Europe: overview.* Retrieved from https://www.openaire.eu/open-science-europe-overview

OpenDOAR. (2019). *OpenDOAR Statistics - v2.sherpa.* Retrieved from: https://v2.sherpa.ac.uk/view/repository_visualisations/1.html

Palcullo, V.E.V., Geromiano, J.F., Superio, D.L. (2019) *Grey literature usage among Filipino aquaculture researchers: A bibliometric analysis of research from 2009 to 2018.* Paper presented at Joint ASFA-UMT Conference Enabling Grey Literature Discovery Benefits Aquatic Sciences, Fisheries and Aquaculture Research, University of Malaysia Terengganu.

Pettman, I. (2019) ASFA trust fund special projects: Open Access university theses. *ASFA Newsletter*, (2), 11-12. http://www.fao.org/3/ca4630en/ca4630en.pdf

Silvoni, G. (2019) *Fisheries Development Project by FAO in the 1960s-1970s: Argentina Case.* Paper presented at Joint ASFA-UMT Conference Enabling Grey Literature Discovery Benefits Aquatic Sciences, Fisheries and Aquaculture Research, University of Malaysia Terengganu.

United Nations (2019). Sustainable Development Goals. Retrieved from https://www.un.org/sustainabledevelopment/sustainable-development-goals/

United Nations (2018) A dialogue on equal, open access and the SDGs. Retrieved from: https://www.un.int/news/dialogue-equal-open-access-and-sdgs

Varley, A. (1995) ASFA: the first twenty years; an outline history of Aquatic Sciences and Fisheries Abstracts, 1971-1990. Retrieved from: http://hdl.handle.net/1834/2656

Wells, P.G. (2014) Managing ocean information in the digital era – Events in Canada open questions about the role of marine science libraries. *Marine Pollution Bulletin*, (83) 1 pp. 1-4, https://doi.org/10.1016/j.marpolbul.2014.04.012

TIB
LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY

"AS AN INFORMATION CENTRE FOR THE DIGITISATION OF SCIENCE AND TECHNOLOGY, OUR OBJECTIVE IS TO SUPPORT RESEARCHERS AT ALL STAGES OF THEIR WORK BY PROVIDING THEM WITH OUR SERVICES."

Professor Dr. Sören Auer

WWW.TIB.EU

ConfIDent – An Open Platform for FAIR Conference Metadata*

Stephanie Hagemann-Wilholt, Christian Hauschke, and Margret Plank
TIB, German National Library of Science and Technology –
Leibniz Information Centre for Science and Technology University Library, Germany

Abstract
Currently, information on scientific events such as conferences is often scattered and not available in the long term. With the project ConfIDent we want to develop a service platform for the quality-driven, collaborative curation of semantically structured metadata of scientific events. It will provide reliable and transparent data and workflows for researchers (organizers, speakers, participants) as well as other stakeholders of scientific events such as university administrations, libraries, sponsors, publishers or specialized societies. The sustainability of the service will not only be obtained a user-centered approach but also by connecting it to existing services enabling data exchange, and by the commitment to the FAIR principles. ConfIDent will reach the current desideratum of long-term findable, open, referenceable and reusable metadata on scientific events.

Introduction

Conferences are a central, in some disciplines indispensable element of scholarly communication. They allow a broad and quick overview about new research topics and areas and present an opportunity to

- network with your community;
- get informal and fast feedback from peers independent from long publication lifecycles, and
- publish articles as conference proceedings.

Information on conferences and their resulting outputs such as proceedings, videos, reports and other formats of documentation can be found on numerous platforms that function as disciplinary and interdisciplinary services. However, there are three major challenges with metadata on conferences:

1. *Availability* of scholarly event data: The data is often scattered over service platforms, temporary websites, newsletters, etc. Further services are used to publish and archive proceedings and their metadata. These services are not linked to each other and not a small size of data gets lost when conference websites disappear after a few months.
2. The second problem is closely connected to the first one and deals with the *quality* of scholarly event data. There exists no uniform standard for conference metadata. The data that is provided by service platforms is often insufficient, very little structured and/or not available in the long-term. This lack of sustainable event metadata makes the tracking of conference activities and their output a very time-consuming task.
   Moreover, non-uniform indexing standards make the disambiguation of conference titles more difficult. In particular, this supports the business models of predatory conference organizers to advertise their events with labels of prominent conferences. Especially for young scholars or scholars from foreign research fields it can be difficult to differentiate between serious and fake conferences. The assessment of the content of conferences and their quality requires the insider knowledge of field experts.
3. The commitment to conferences, e.g. by organizing them, presenting there, accomplishing review tasks, can take up a considerable part of the work of researchers without any *acknowledgement* of these activities as research output. The academic system still only rewards publications as evidence of scientific activity and tries to further condense them with the help of singular indicators.

* First published in the GL21 Conference Proceedings, February 2020.

An analysis of 27 conference platforms has shown that considerable deficits of existing services lie in the long-term findability, availability and accessibility of event metadata and content information. We chose platforms that are frequently used, highly advanced and/or have a broad community approach. The majority of platforms do not use any persistent identifiers (PIDs) and metadata is often inadequate: In some cases they do not even provide core metadata such as the full title of a conference or a working URL for upcoming events. Often, events are treated as singular entities and are not related to superordinate event series. Links between contributors, contributions, proceedings and affiliations are an exception; most content information (e.g., abstracts, organizers, speakers, research topics) are either unstructured or can only be found on temporary event websites. Best practices can be seen in community driven approaches which seem to attract more users and to encourage better curation of contents.

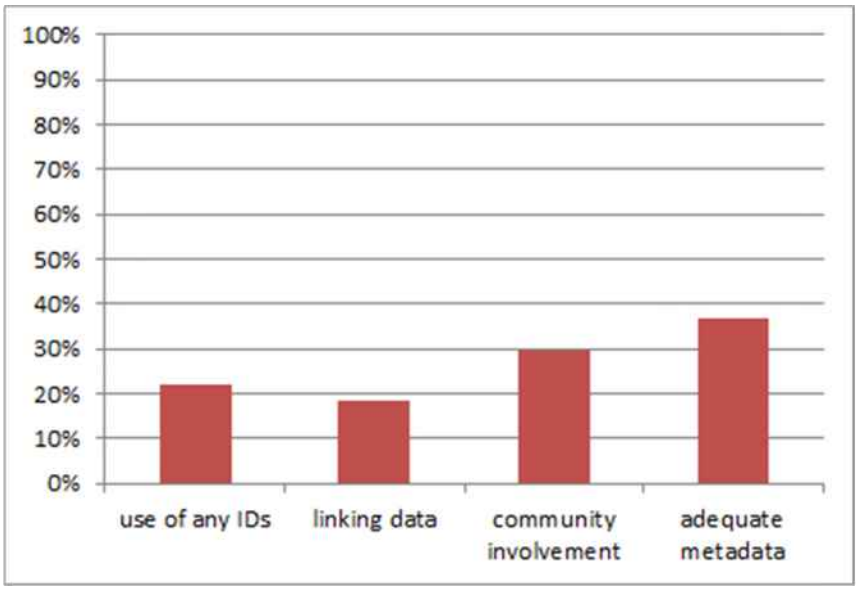


Figure 1: Analysis of 27 conference platforms services and their key components.

The ConfIDent Project

Objectives

In December 2019 ConfIDent started, a joined project of Technische Informationsbibliothek (TIB) Hannover and RWTH Aachen University which is funded by German Research Foundation (DFG – Deutsche Forschungsgemeinschaft).[1] The project aims to establish a service that provides data on scientific events which enables researchers to find, promote and archive information. It is designed as a pilot project developing a prototype for two research communities – a) computer and data science and b) research on transport and mobility – considering the specific relevance of scientific events within these communities. A user-centred approach supports the development of a service that reflects the information needs of the scientific communities. In addition, we want to empower the users to curate the data on their own in order to use their field expertise. A standardisation of event metadata will not only allow data exchange with existing databases but will also foster the assignment of PIDs to improve the quality of metadata. This is a prerequisite to meet the FAIR data requirements and make conference metadata findable, accessible, interoperable and re-usable in the long-term.[2]

User-centred approach

In order to develop and operate a service platform for event related scholarly metadata that is geared to community needs of specific target groups a user-centred design approach[3] is applied.

[1] Project Website ConfIDent: https://projects.tib.eu/en/confident (access: 20/01/14).
[2] FORCE11 – FAIR Data Principles: https://www.force11.org/group/fairgroup/fairprinciples (access: 20/01/14).
[3] ISO-Standard 9241-210:2019: Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems: https://www.iso.org/standard/77520.html (access: 20/01/14).

This ensures that the platform will be easy and intuitive to use and that the user can effectively and efficiently achieve the desired result during the interaction with the system. First of all, information is collected with potential users from the target groups regarding their typical usage behaviour, tasks, needs and expectations regarding the platform. In particular, the domain-specific characteristics of conferences will be taken into account. The target groups will be surveyed by typical user centred design methods such as brainstorming, task analysis, user stories, personas, focus groups and interviews. The analysis results of the user requirements form the basis for the identification of the functionalities and the visualisation of the platform. In an iterative approach potential users from the identified user groups will work together with the developers already in the early phases of specification. Two iterations of evaluations are planned during the development process of the platform. In that way errors and difficulties of the different user groups and stakeholders can be identified at early stages of the process. Central points of the evaluation are the effectiveness, the usability, and the user experience of the platform. Measuring and testing these aspects is achieved by a combination of different methods like task-based usability testing combined with eye tracking, thinking aloud, user interviews and questionnaires. We plan on a first prototyped version after six months of development. A revised and optimized version will be created in the first project year.

In order to become a long-lasting, reliable and accepted platform for science conferences and other scientific events and to reach a critical mass of users the development of a scholarly and a technical community around the ConfIDent project is an important objective. In order to accomplish this, we will establish connections of the ConfIDent platform to existing services including library catalogues, research information systems such as VIVO,[4] publishing platforms like Open Journal Systems e.g. Copernicus Publications[5] or the TIB AV-Portal[6] for the provision of conference recordings. By connecting the resources of various existing services, ConfIDent ensures permanent links between different resource types (such as proceedings, recordings, contributor profiles, organizers etc.). The provision of conference metadata and the assignment of persistent identifiers allow both researchers and infrastructure providers an improved disambiguation and quality assessment of scientific events.

Based on the requirements of users and researchers from the target communities as well as their usage behaviour, various business models will be developed within the framework of the project. These will be discussed and further developed in user workshops. Moreover, the project partners will raise awareness in the target communities for the benefits of ConfIDent by attending events, sending announcements and invitations. In particular, this effort will be supported by the German Informatics Society (GI)[7] in computer science.

The communities will also be engaged into contributing to the ConfIDent software and platform by inviting research communities to use the ConfIDent platform as a test bed for their developments and tools such as recommender systems, graph partitioning, and clustering methods.

Based on the result of the user needs analysis and in close cooperation with the communities, guidelines for potential content deliverers will be developed, which clearly define the portal's profile. Further user workshops, webinars and training materials such as explanatory videos will be developed to support the communities when using the portal.

Metadata quality, PIDs, and scientometric indicators

One of the core aspects of the project is to define a framework for high quality of metadata and content. Rich metadata helps to disambiguate event data and provides context information on events. PIDs strongly support the standardization of metadata, and are a prerequisite for FAIR data including long-term availability of information.[8]

[4] Conlon et al. (2019).
[5] Copernicus Publications: https://publications.copernicus.org (access: 20/01/14).
[6] TIB AV-Portal: https://av.tib.eu (access: 20/01/14).
[7] Gesellschaft für Informatik: https://gi.de (access: 20/01/14).
[8] Demeranville (2018).

ConfIDent maps existing machine readable metadata schemes as used e.g. by DataCite, Crossref, ORCID or ROR for PID registration as well as schema.org., the RDA Framework of the Integrated Authority File (GND - Gemeinsame Normdatei) used by the German National Library and the cataloguing system of the Common Library Network (GBV - Gemeinsamer Verbundkatalog).[9] In this way, interoperability with existing services should be achieved. Furthermore, we are actively engaged in the international Working Group initiated by DataCite and Crossref to develop a Conference PID.[10] This PID will provide a response to the specific metadata requirements for events as a resource type and will support a standardization of conference information which is still a desideratum. Standardization will help to support a quality assessment approach to conference information both on the metadata and the content level. ConfIDent will have a tiered metadata concept and differentiate between mandatory, recommended and optional fields. On the one hand, this modular approach offers a minimum, generic set of metadata that is necessary to identify events unambiguously; on the other hand it allows subject-specific adjustments as the event related information needs may differ from research community to research community. Content quality criteria for conferences can to a certain extent be represented by metadata as indicators. However, it must always be considered that

1. each indicator has a limited value in itself; and
2. quality criteria for conferences are extremely dependent on the professional culture of each research field or community.

Therefore, we want to define the metadata requirements together with experts from the pilot communities in order to take into account their information and quality requirements. Rich metadata offers more options to assess the quality of an event, but it also means more effort for data providers and curators. These users should also be recruited from the scientific communities, but we want to balance information requirements with the workload for the individual user.

Technical base

Building upon the experiences of both project partners with OpenResearch.org[11] (OR) – an experimental platform for the analysis of research information on events, papers, projects and other entities – the project will start with a prototype based on the open source software Semantic MediaWiki (SMW)[12] fostering openness and extensibility. The SMW prototype will be customized to metadata and user requirements identified in an iterative process as described by the user-centred approach. The software allows collaborative data curation and options for a high degree of transparency to display the provenance of data. Nevertheless, this collaborative approach challenges the objective of high quality metadata. ConfIDent aims to provide a mature rights and roles managements as well as guidance for data ingest and curation to allow easy use but prevent misuse. We want to provide an open platform with interfaces to allow data exchange with existing services and an attractive platform for individual users and their information needs.

Extensional use of data

The use of scientometric indicators to measure the impact of scientific output is heavily discussed, with some key documents defining the outlines of what is called altmetrics (alternative metrics).[13] These metrics track the use of research outcomes online, such as on social media, news sites, blogs and policy papers. Compared to conventional metrics such as citation counts, altmetrics provide

[9] Crossref Metadata Schema 4.4.2 on conferences: http://data.crossref.org/reports/help/schema_doc/4.4.2/schema_4_4_2.html#conference; DataCite Metadata Schema 4.3: https://schema.datacite.org/meta/kernel-4.3/doc/DataCite-MetadataKernel_v4.3.pdf; GND Erfassungshilfen für Körperschaften und Konferenzen: https://wiki.dnb.de/pages/viewpage.action?pageId=90411359; ORCID Metadata Schema 3.0: https://github.com/ORCID/ORCID-Source/tree/master/orcid-model/src/main/resources/record_3.0; schema.org event: https://schema.org/Event (access: 20/01/14).

[10] Birukou (2018).

[11] Vahdati et al. (2016).

[12] Semantic MediaWiki: https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki (access: 20/01/14).

[13] Fraumann (2018); Tunger et al. (2018).

among others a time advantage and the possibility to include mentions outside the scientific communities.[14]

Conclusion

The central objective of the project is to develop a sustainable service platform that provides reliable data on scientific events. ConfIDent will not only facilitate quality assessment of data with regard to a wide range of criteria and stakeholders' perspectives, taking into account broad context information. The platform will also foster a cultural change in science by providing a higher visibility of scientific events as an independent achievement of research beyond counting article citations and by promoting their impact. We see the initial community oriented approach as starting point for the development of a generic service that serves the scientific community as a whole. The service is supposed to be connectable to numerous projects and initiatives that aim to better capture the heterogeneity of scientific outputs and making them accessible.[15]

Literature

Auer, Sören; Blümel, Ina; Ewerth, Ralph; Garatziogianni, Alexandra; Heller, Lambert; Hoppe, Anett; Kasprzik, Anna; Koepler, Oliver; Nejdl, Wolfgang; Plank, Margret (2018): Towards An Open Research Knowledge Graph (Zenodo). DOI: https://doi.org/10.5281/zenodo.1157185.

Birukou, Aliaksandr (2018): PIDs for conferences - your comments are welcome! (DataCite Blog). DOI: https://doi.org/10.5438/prq9-ed42.

Collins, Sandra; Genova, Françoise; Harrower, Natalie; Hodson, Simon; Jones, Sarah; Laaksonen, Leif et al. (2018): Turning FAIR into reality. Edited by European Commission Expert Group on FAIR Data. Directorate-General for Research and Innovation. DOI: https://doi.org/10.2777/1524.

Conlon, Michael; Woods, Andrew; Triggs, Graham; O'Flinn, Ralph; Javed, Muhammad; Blake, Jim et al. (2019): VIVO: a system for research discovery. In JOSS 4 (39), p. 1182. DOI: https://doi.org/10.21105/joss.01182.

Demeranville, Tom (2018): Building a robust research infrastructure, one PID at a time (ORCID Blog). Available online at https://orcid.org/blog/2018/08/08/building-robust-research-infrastructure-one-pid-time.

Fraumann, Grischa (2018): The Values and Limits of Altmetrics. In New Directions for Institutional Research 2018 (178), pp. 53–69. DOI: https://doi.org/10.1002/ir.20267.

Hauschke, Christian; Cartellieri, Simone; Heller, Lambert (2018): Reference implementation for open scientometric indicators (ROSI). In RIO 4, p. 59. DOI: https://doi.org/10.3897/rio.4.e31656.

Hicks, Diana; Wouters, Paul; Waltman, Ludo; Rijcke, Sarah de; Rafols, Ismael (2015): Bibliometrics: The Leiden Manifesto for research metrics. In Nature 520 (7548), pp. 429–431. DOI: https://doi.org/10.1038/520429a.

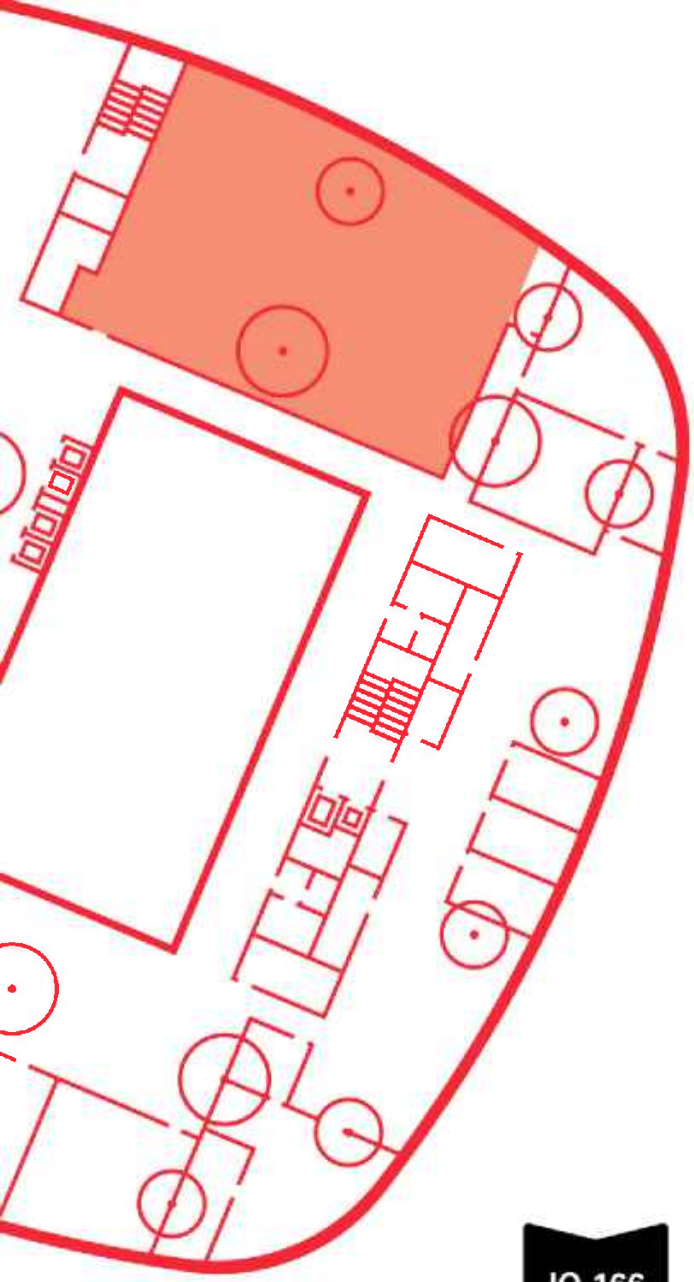
Tunger, Dirk; Clermont, Marcel; Meier, Andreas (2018): Altmetrics: State of the Art and a Look into the Future. In Mari Jibu, Yoshiyuki Osabe (Eds.): Scientometrics: InTech. DOI: https://doi.org/10.5772/intechopen.76874.

Vahdati, Sahar; Arndt, Natanael; Auer, Sören; Lange, Christoph (2016): OpenResearch: Collaborative Management of Scholarly Communication Metadata. In Eva Blomqvist, Paolo Ciancarini, Francesco Poggi, Fabio Vitali (Eds.): Knowledge Engineering and Knowledge Management. 20th International Conference, EKAW 2016, Bologna, Italy, November 19-23, 2016, Proceedings, vol. 10024. Cham, s.l.: Springer International Publishing (Lecture Notes in Computer Science, 10024), pp. 778–793. DOI: https://doi.org/10.1007/978-3-319-49004-5_50.

Wilsdon, James; Bar-Ilan, Judit; Frodeman, Robert; Lex, Elisabeth; Peters, Isabella; Wouters, Paul (2017): Next-Generation Metrics: Responsible Metrics and Evaluation for Open Science. Report of the European Commission Expert Group on Altmetrics. Handle: http://hdl.handle.net/1887/58254

[14] Hauschke et al. (2018); Hicks et al. (2015); Wilsdon et al. (2017).

[15] Auer et al. (2018); Collins et al. (2018); FORCE11 – The Fair Data Principles: https://www.force11.org/group/fairgroup/fairprinciples; PID Graph at FREYA Project: https://www.project-freya.eu/en/about/mission (access: 20/01/14).



National Library of Technology

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

For detailed information on the National Library of Technology visit **http://www.techlib.cz/en/**

95 mm

As corresponds to its statutes, NTK manages – – among others – the project of building the **National Repository of Grey Literature**. The project aims at gathering metadata and possibly full texts of grey documents in the fields of education, science and research. The NTK supports an education in the field of grey literature through annual seminars in the Czech Republic.

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

NTK Univers

Národní technická knihovna
National Library of Technology

NUŠL

COVID-19 and Work from Home: Digital Transformation of the Workforce

Dobrica Savić, Austria

Abstract
The outbreak of COVID-19 around the globe forced businesses to innovate and change the way they conduct their work. Offices have become less important and work from home has suddenly become mandatory. This sudden need for work from home is driving the digital transformation of the workforce and the evolution of the work environment at an unprecedented speed. Mass adoption of telecommuting has become a vital business change since the outbreak of the virus. This paper looks at this unprecedented impact of coronavirus pandemic on sudden demand for work from home and the subsequent push for the digital transformation of the workforce.
Keywords: COVID-19, Coronavirus, Digital transformation, Workforce

Introduction
In the past few years, digital transformation has been among the most discussed topics of business leaders and management gurus. Many organizations, while understanding the importance of digital transformation and giving thought to this new trend, have been slow on the uptake to change the status quo, cause potential disruption, and adopt inevitable transformation. Effort has been placed on smaller projects, minor changes, learning from others, and waiting for a more suitable time to make radical change. With the onset of 2020 and the outbreak of COVID-19, organizations found themselves caught off guard and ill-prepared for the new normal brought about by the coronavirus reality. "With millions of people retreating to the safety of the online world for their news, entertainment, education, communication and remote work — the imperative of digital transformation has gone from important to absolutely critical."[1]

COVID-19
At the end of December 2019, Chinese health officials informed the World Health Organization (WHO) about a cluster of patients with a mysterious pneumonia. On 30 January 2020, WHO declared the COVID-19 outbreak a public health emergency of international concern. *Pandemic* suddenly became a household word inundating our daily lives, and the definition — an infectious disease where we see significant and ongoing person-to-person spread in multiple countries around the world at the same time[2] — familiar to all.

Within days of the outbreak, work from home (WFH), until then practiced sporadically by companies and organizations, became mandatory — a question of physical and financial survival. To prevent the spread of the virus and protect workers, governments around the world instructed employers to close their offices and enable employees to work from home — to *telecommute*. It is estimated that more than four out of five people (81 percent) in the global workforce of 3.3 billion are being affected by full or partial workplace closures.[3]

Many, employers and employees alike, had limited experience with WFH on such a large scale before the outbreak. Companies had been comfortable following the standards of office-based work and employees were comfortable with the dichotomy of work and home life. Suddenly companies were faced with the very real possibility of losing revenue and jobs. Combined with the fear and uncertainty of the virus, this disruption was huge, stressful, and painful. WFH was looked on as a potential solution to relieve the pain.

Work from home
Work from home, a phrase commonly used since the onset of COVID-19, can be defined generically as employees working outside company offices. It includes four basic characteristics: (1) a person who is an employee of a company or a staff member of an organization; (2) actual work engagement with a company or an organization on specific tasks; (3) work being performed outside the company's physical premises; and (4) telecommunication with the employer.

It is worth mentioning that two other popular terms, often used synonymously to describe work from home, are *telecommuting* and *remote work*. The figure below from Google Trends of searches

made worldwide in the past three months for *telecommuting* and *remote work*, indicates that *remote work* was used more frequently.

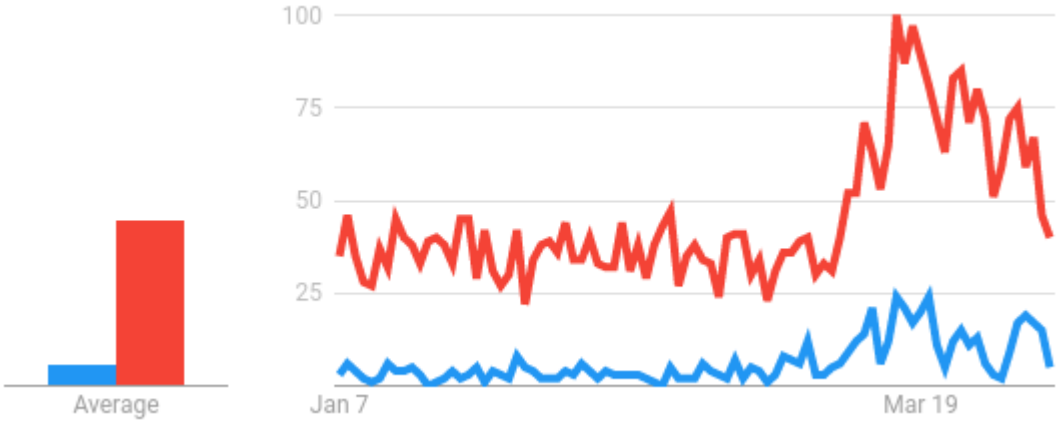


Figure 1: Google Trends — frequency of searches for *telecommuting* (blue) vs *remote work* (red) from 7 January to 6 April 2020

Not only are searches using *remote work* more frequent, but the availability of information resources mentioning remote work is also greater. There are 17.2 million pages with the term *remote work* compared to 13.9 million web pages with the term *telecommuting.* Conversely, Google Scholar offers 51,000 articles on telecommuting and 13,700 on remote work.

Often used interchangeably, *remote work* and *telecommuting* have subtle differences in meaning, indicating two somewhat different approaches to the concept of working from home. For the sake of clarity, it is worth defining them in more detail.

Telecommuting is a work arrangement in which the employee works outside the office. This often means working from home or from locations close to home, such as coffee shops, libraries, or co-working spaces.[4] Rather than commuting to the office, employees 'travel' using IT tools, keeping in touch with coworkers and employers via mobile devices, telephones, online chats, video conferences, collaboration platforms, and email.

Allen, Golden, and Shockley (2015) conducted a comprehensive review of a wide range of telecommuting studies and found that most definitions of telecommuting have two things in common: working from a location other than the traditional office, and using technology to perform work-related tasks. They also identified three areas in which the definitions differ: (1) the extent of telecommuting (e.g., once a month, once a week, full-time), (2) the type of employment relationship (e.g., staff member, independent contractor or consultant, self-employed), and (3) the location of remote work (e.g., home, satellite office, coffee shop) (Allen et al., 2015). [5]

Benefits of telecommuting for both employees and employers, besides saving money, include increased job satisfaction and productivity, greater flexibility, reduced office costs and requirements, increased staff retention, improved employee work-life balance, keeping older generations in the workforce, and environmental benefits.

However, disadvantages also exist such as sociological and psychological challenges resulting from isolation, long work hours, and lack of separation between work and home. Telecommuting requires self-discipline and dependence on personal IT tools, communication, and other resources. There are also the costs involved for organizations when transitioning to new work methods and training, along with employees' lack of both commitment and identification with organizational culture and values. The increased risk of privacy and security issues should not be underestimated. Some authors also mention the increased danger society faces by creating detached individuals.[6] [7]

Remote work, another type of work from home, is seemingly synonymous with telecommuting. However, there is a slight difference between these two terms. While telecommuting means working outside the office, usually from home, remote work implies that the employee lives outside the vicinity of the organization's main headquarters or office. This geographical distinction may seem inconsequential, but in fact, it demands change in the management and engagement of the workforce. Managers need to adopt different communication and management styles and

make additional efforts to properly lead and ensure the required level of productivity of a remote workforce. Different labor legislation, financial obligations, cultural backgrounds, time-zones, scheduling, and expectations are just some of the dissimilarities between remote work and telecommuting. Often, remote workers are freelancers and independent contractors who spend their time outside the traditional office settings.

The benefits of a diverse globally distributed workforce boil down to the large pool of specialty professions and the financial gains brought about by different standards of living and local pay. Remotely recruited employees often have considerable financial benefits such as higher salaries, better working conditions, and international exposure.

The disadvantages mentioned above for telecommuting are more or less the same for remote work. However, disadvantages unique to remote workers include their emotional well-being, limited career development opportunities, working outside regular local work hours due to different time zones, professional or even geographical bias, and emotional resentment of coworkers and local colleagues.

Digital transformation

The hard reality of the impact COVID-19 has had on the economy and people's livelihoods has brought the concept of digital transformation into focus. This has especially been the case for the hard-hit workforce. Because WFH has become inevitable, new work models have had to be quickly developed and deployed. Terms such as telecommuting, teleworking, working from home, working at home, working remotely, virtual work, e-work, e-commuting, mobile work, flexible workplace, digital nomads, and freelancing have all been used to describe the current modes of work and to jump start digital transformation of the workforce.

Digital transformation is about doing things differently — creating a completely new business model by using modern information and computer technologies. Digital transformation leverages existing knowledge to profoundly change the essence of the organization — its culture, management strategy, technological mix, and operational setup. It places the customer at the center of all its decisions and actions.[8]

The ideas and solutions of Industry 4.0 — digital transformation — have quickly become, for many, a panacea to the COVID-19 disruption. Changing business models and work procedures, maximizing the use of modern information technology (IT), requiring adjustments to organizational culture and behavior, and modifying the expectations and roles of the workforce have become the new rules of the game.

The impact of COVID-19 on the workforce is visible on multiple levels. This includes a change in the nature of work, its variety, volume, velocity, and value. Digital transformation is more than just the implementation of a new technology. It requires the adoption of a "digital workforce mindset"[9]. A digital mindset involves a deep understanding that the power of technology can democratize, scale and speed up every form of action and interaction. The main characteristics of a digital mindset are: abundance, growth, agility, comfort with ambiguity, an explorer's mind, collaboration, and embracing diversity.

Digital transformation of the workforce requirements include[10]:
- Digital literacy, technical knowledge
- Lifelong micro learning and personal development
- Engagement
- Mobile force and remote work
- Generation gap
- Digital ethics

Even before the coronavirus outbreak, the World Economic Forum (2018) estimated that by 2022 over 50% of all employees would require significant reskilling and upskilling. This will be a huge task for HR and other managers, especially since 85% of 2030 jobs don't yet exist.[11]

Conclusion

As the ripple of COVID-19 careens around the globe, we are being forced to innovate and change the way we work and live[12]. Offices have become less important and work from home has suddenly become mandatory. This sudden need for work from home is driving the digital transformation of the workforce and the evolution of the work environment at an unprecedented speed. Mass adoption of telecommuting has become a vital business change since the outbreak of the virus. And this change is here to stay.

In a matter of days, organizations have been required to improve their capabilities for long-distance collaboration. Video conferencing, online purchasing, special deliveries, telemedicine, e-learning, electronic trading, online marketing, video streaming, and many other IT enabled processes have undergone virtual transformation, replacing traditional work practices. Digital transformation covers a wide spectrum, including maximizing the use of modern information technology. Because of COVID-19, it has gained importance and been widely recognized and accepted by both employers and employees.

The recent transformation of the workforce is a crucial step forward for digital transformation. Organizations that have enhanced their IT capabilities and remotely engaged their employees are in a much better position to not only survive these unprecedented circumstances, but to overcome the short and long term challenges that will inevitably follow.

References

1 The Coronavirus and Public Service Media: Why digital transformation matters now more than ever. Sasha Scott. EDU Blog. 31 March 2020. https://bit.ly/3aAZUIb

2 Coronavirus: What is a pandemic and why use the term now? BBC, 11 March 2020. https://www.bbc.com/news/health-51358459

3 ILO: COVID-19 causes devastating losses in working hours and employment. 7 April 2020. https://bit.ly/2XjoxFu

4 Doyle, A. (2020). What is telecommuting? The balance careers: Basics – Glossary. https://bit.ly/2y9MuVd

5 Allen, T. D., Golden, T. D., & Shockley, K. M. (2015). How effective is telecommuting? Assessing the status of our scientific findings. Psychological Science in the Public Interest: A Journal of the American Psychological Society, 16(2), 40–68. https://doi.org/10.1177/1529100615593273

6 Harpaz, Itzhak (2002). Advantages and Disadvantages of Telecommuting for the Individual, Organization and Society. Work Study 51(2):74-80, April 2002. https://bit.ly/2VmJJb2

7 Vaganay, Arnaud; Canónico, Esther; Courtin, Emilie (2016). Challenges of work-life balance faced by working families. European Commission, Evidence Review. May 2016. https://bit.ly/2RtJ0Ul

8 Savić, Dobrica (2019). From Digitization, through Digitalization, to Digital Transformation. 43/2019. 36-39. https://bit.ly/3aO3Gy3

9 Chattopadhyay, Sahana (2016). 7 Characteristics of a Digital Mindset. People Matters. https://bit.ly/2mzNplZ

10 Savic, Dobrica (2020). Digital Transformation and Grey Literature Professionals. Grey Journal, February 2020 16(Special Winter Issue):11-17. https://bit.ly/2XeMA8u

11 DELL Technologies (2018). Realizing 2030: A Divided Vision of the Future. https://bit.ly/2FvF1yi

12 Marr, Bernard (2020). 9 Future Predictions For A Post-Coronavirus World. Forbes, April 3 2020. https://bit.ly/2JNWPsx

E-LIS between old and new forms of Grey Literature encompasses new forms of relationship between librarians in the different country*

Antonella De Robbio, Imma Subirats, E-LIS ePrints for Library and Information Science, Spain
Fernanda Peset, Polytechnic University of Valencia, Spain

Abstract

E-LIS ePrints for Library and Information Science is an international digital repository for Library and Information Science (LIS) Open Archives Initiative compliant. This 21th International Conference is focused on open science this year and "Grey literature by definition seeks to make publications produced on all levels of government, academics, and business openly accessible different from those controlled by commercial publishing." (http://www.textrelease.com/gl21conference.html) Thematic open access repositories as arXiv or E-LIS have years hosting this kind of academic materiales. E-LIS hosts documents in 22 document types (plus dataset recently added) including those belonging to the traditional gray literature world as Preprints, Thesis, Technical and dept. Reports, and those encompassing new forms of Grey Literature as Data and Datasets. These old and new forms of grey Literature constitute 40% of the whole content of the repository (working papers presented to conference, congresses and different events, presentation in PPT, tutorial and learning material for almost, report, thesis, …)

E-LIS is completely built with open source softwares is hosted by University Federico II.

After 16 years from its take-off in 2003, the disciplinary repository contains 22,000 open access articles, in 27 different languages. All the continents are now represented, with a distribution that actively involves over 60 different countries. There are over a million annual discharges requested by users all over the world, numerous from the United States but also from China and South America. The metadata of the deposited works (one hundred per month on average) are validated by the editorial staff divided by geographical areas. The pivot on which the editorial work gravitates is the classification by country which has conferred an international aspect to the archive, aligning it with the organizational model, possibly thanks to the voluntary collaboration of about 80 professionals, including editors and technicians. The collaboration with countries ignored for years by the librarianship tradition is what has made E-LIS particularly innovative on issues previously poorly represented or considered on the margins. Topics that reflect a "different" cultural approach in E-LIS finds space, generating an intellectual growth with respect to the comparison between identity and otherness, in particular respects to the presence of Grey Literature. The studies on the bibliometrics of Indian colleagues, the collaboration with Cuban librarians at the time of the US embargo, the request to include the Maori language by New Zealand colleagues, the emergence of contents from the East of the world, the ferment of the South American jobs for Open Access are just some of the inclusiveness traits of the multicultural character of E-LIS, one of the reasons for its success. Its organizational structure of international scope makes it a model for the construction of open digital libraries, exportable to other communities. Its technical structure, linked to the OAI context and the accompanying innovative tools, provides useful services for the scientific communication circuit: analysis of log files for the production of statistics by author and for single work. Recently the connection to the Zenodo server provides the management of research data in a transparent and integrated way in a new mode to communicate LIS disciplines.

E-LIS in the Open Access world

E-LIS is an international digital repository for Library and Information Science (LIS), including Communication. Created in 2003 and hosted by Università di Napoli Federico II in Italy, in a few years has been indeed as the largest international open repository in the field of library and information science. The strength of our archive is that all the work is completely based on voluntary work. It has grown thanks to a team of 80 volunteer editors, LIS professionals as technicians, librarians and information specialists. The editorial team is formed by 67 editors

* First published in the GL21 Conference Proceedings, February 2020.

coming by different countries plus a team of technicians and the Administrative Board which decides policies and rules.
After 16 years from its take-off, the disciplinary repository contains 22.000 open access contents in 27 different languages.

Illustration 1.

The library and information world is highly integrated with the areas of computing science and technology and it was felt that the LIS discipline should set an example to other communities by providing a state-of-the-art model for the OA movement and digital libraries, particularly in relation to the open archive model, within which E-LIS is a disciplinary repository. The extension of the OA concept to LIS works and the dissemination of material within the LIS community will contribute to the development of an international LIS network; E-LIS is mutually beneficial. For library and information science professionals Meta creation is costly and the growing trend of authors to self-archive in the OAI framework is proving an effective way to reduce some of those costs. For librarians as authors, archiving their work in E-LIS gives them an increased understanding of the process of self-archiving and the E-LIS archive ensures data preservation and a wide data visibility in addition to facilitating active participation in the international community of library and information science professionals.

Main focus of E-LIS is its organizational model and on the strategic issues correlated with Open Access (OA). Some of the challenges and opportunities consequent on a global vision for the Library and Information Science (LIS) field, envisages papers coming from all over the world and which gives E-LIS the impetus and motivation to stimulate participation in the venture and to further develop international research activities.

According to Foster Consortium, an institution of the European Union dedicated to the diffusion of Open Access idea, Open Science is "the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society". E-LIS is inserted in the Open Science context, uses Open Source software ePrints and is OAI compliant. All contents inside the repository are all full-text and open access.

Its main focus, however, is on the E-LIS organizational model and on the strategic issues correlated with Open Access (OA). It also delineates some of the challenges and opportunities consequent on a global vision for the Library and Information Science (LIS) field which envisages papers coming from all over the world and which gives E-LIS the impetus and motivation to stimulate participation in the venture and to further develop international research activities.

Metadata and Dataset management as quality guarantee

E-LIS puts a great attention on metadata quality**.** Cultural and memory institutions have a long tradition of setting up, publishing, and sharing vast amounts of metadata, such as library catalogues and archival finding, providing inventories of books and documents with detailed descriptions of individual items using many different formats and approaches and the editorial's team of E-LIS is mainly constitutes by librarians. Because librarians are so involved in open access advocacy, e-LIS is a key to encourage open access for all repositories, by giving librarians the experience they need to speak with confidence when talking with researchers and open access archives, and the experience to provide the best possible assistance to self-archiving faculty.

E-LIS Metadata are set on the basis of a 23 document typologies and then checked by the international editors committee in accordance with editorial guidelines. One of the first step of the European vision was to establish guiding principles for individual datasets. Since 2014, if possible, Data should be FAIR, that is, Findable, Accessible, Interoperable and Reusable, and policies, tools, infrastructures have to interoperate to fulfill these principles. One on the 23 typology of metadata concerns "dataset" which can be described and deposited as a item itself inside the repository.

When we talk about research data we mean all the recorded information necessary to support or validate a research project.We have to take into consideration also digital objects, simple or complex**.** If we want to manage digital objects in an Open Access vision, they must be available in structured collections or stored in a computer system. Data can be divided into general categories, according to the way they have been collected or elaborated, and these are the most common types.

Its technical structure, linked to the OAI context and the accompanying innovative tools, provides useful services for the scientific communication circuit: analysis of log files for the production of statistics by author and for single work. Recently the connection to the Zenodo server provides the management of research data in a transparent and integrated way in a new mode to communicate LIS disciplines. For documents that have a data set as useful kit for the validation of the LIS research, we have enclosed a new field as addition into papers metadata set named LINK TO RESEARCH DATA. Such link connect to Zenodo data repository a tool for Big Data management and extended Digital Library capabilities for Open Data developed by CERN in Geneve to support European research programmes. The name Zenodo is derived from *Zenodotus*, the first librarian of the Ancient Library of Alexandria and father of the first recorded use of metadata, a landmark in library history.

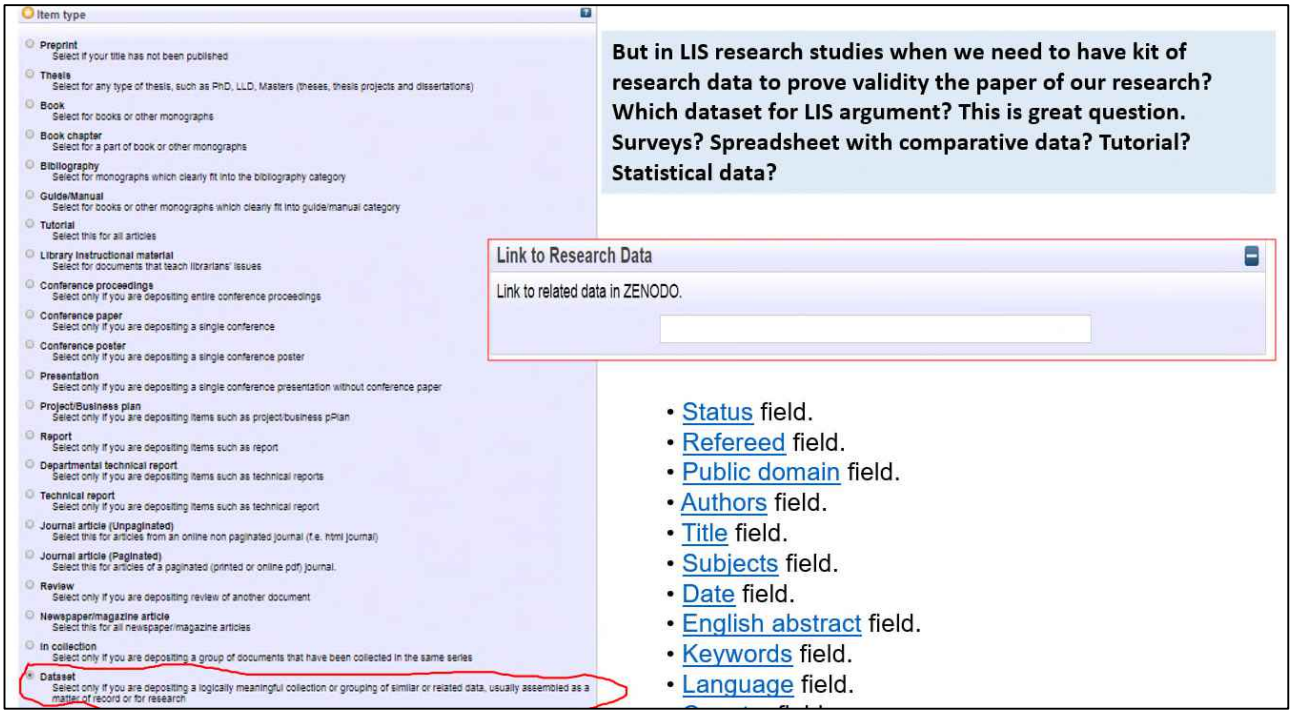


Illustration 2.

The possibility to get statistics give us the dimension of the real interest by users coming by all the world. There are over a million annual download requested by users all over the world, numerous from the United States but also from China and South America. This data means that the repository is live and well knowing not only by LIS communities, also it is a referral point for research of papers published and Grey Literature.

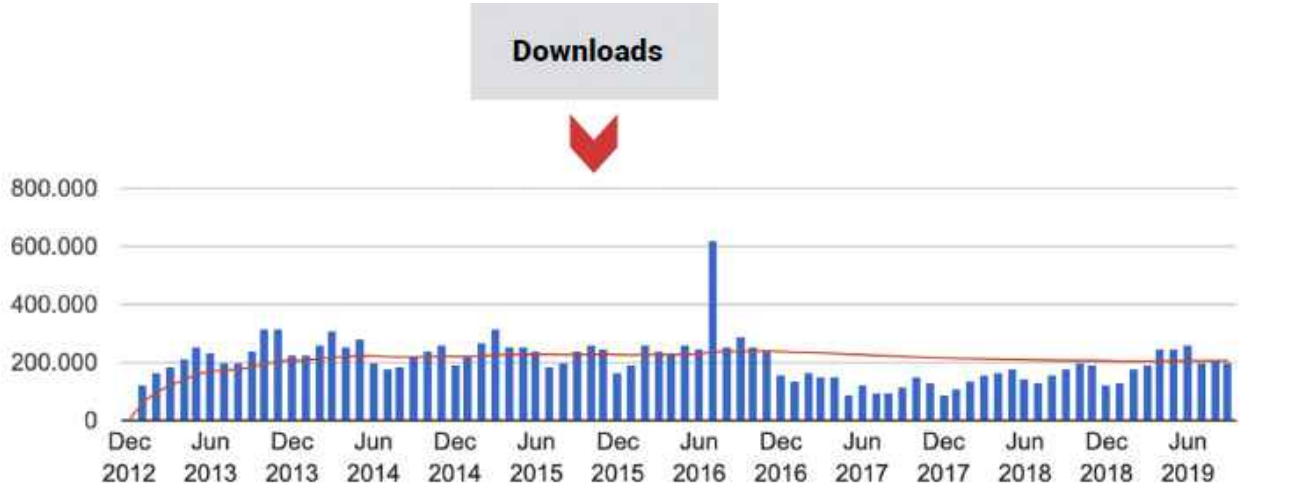


Illustration 3.

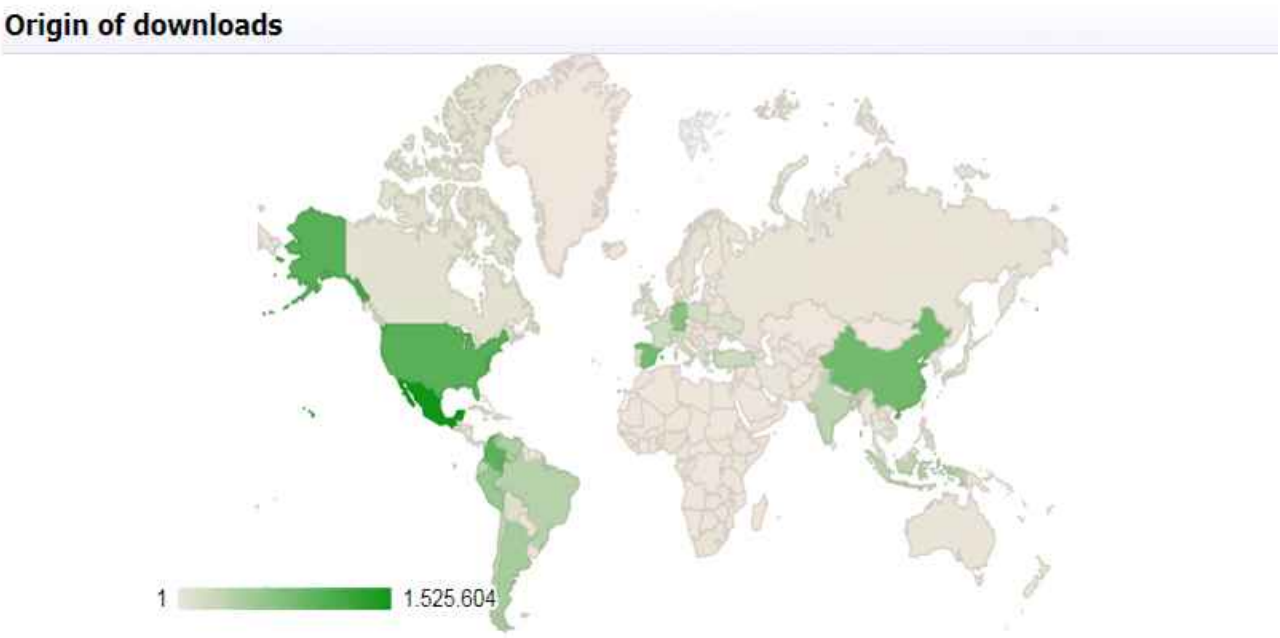


Illustration 4.

The Grey world inside E-LIS: a wide-ranged scale of grey

One of the most recent developments in the field of Library and Information Science (LIS) is the trend towards digital libraries and self-archiving. Self-archiving can be defined as the deposit of a digital document in a public, free-access repository, for example, an e-print archive. An e-print archive is a collection of digital research documents such as articles, book chapters, conference papers and data sets. E-prints are the digital texts of peer-reviewed research articles, before and after refereeing. Before refereeing and publication, the draft is called a "preprint".

Another consideration is that E-LIS accept 35 different formats of document, but the PDF is the prince format

The refereed, accepted, final draft is called a "postprint". The term e-prints include both preprints and post prints.

Documents deposited – almost 100 per months - may include preprints, postprints, conference papers, conference posters, presentations, books, book chapters, technical reports/departmental working papers, theses, and newspaper and magazine articles.

Illustration 5.

The E-LIS submission policy states that the archive accepts any scientific or technical document, published or unpublished, on librarianship, information science and technology or related activities.

In this context, categories for different types of material have been created with respective sets of metadata.

In other terms E-LIS hosts documents in 22 document types (plus dataset recently added) including those belonging to the traditional gray literature world as Preprints, Thesis, Technical and dept. Reports, and those encompassing new forms of Grey Literature as Data and Datasets. These old and new forms of Grey Literature constitute 40% of the whole content of the repository: preprint 5%, thesis 4%, report 2%, another 20% concerning working papers presented to conference, congresses and different events. In addition are depositing also presentation in PPT, tutorial and learning material for almost 9%.

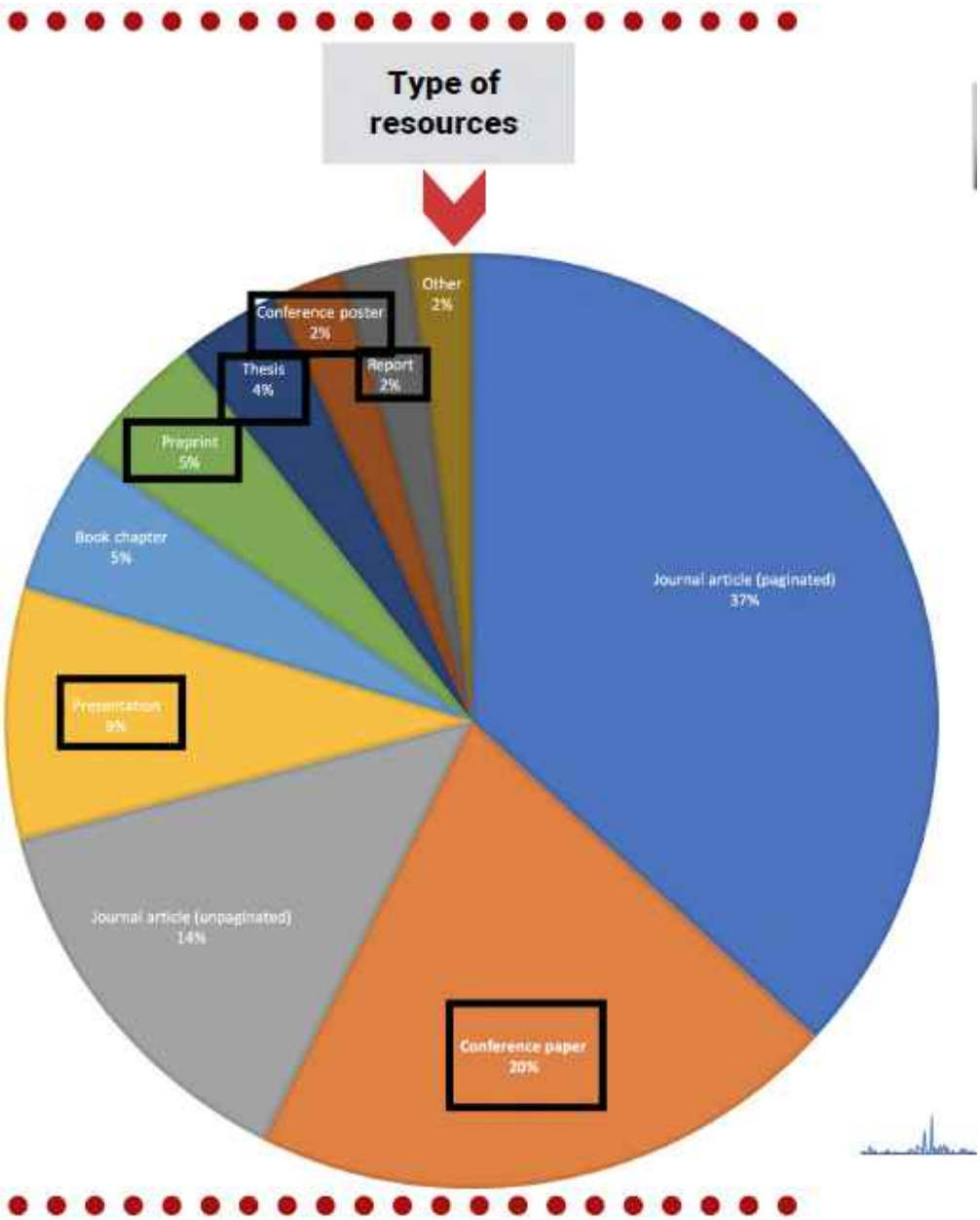


Illustration 6.

Here an example of a metadata for a preprint

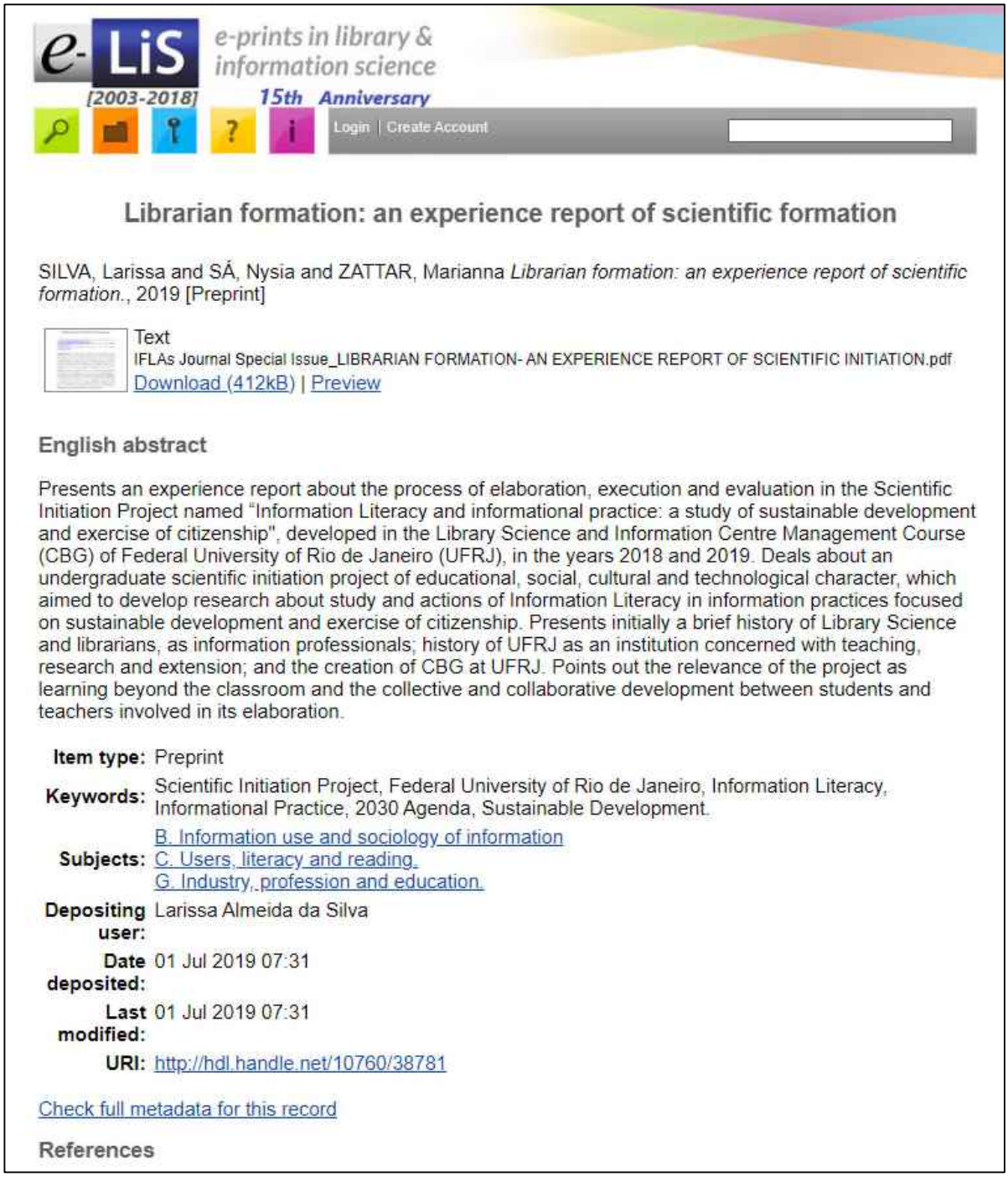


Illustration 7.

E-LIS Policies and its organisational model

E-LIS is driven and directed by its policies, which determine its identity, quality and direction. Merely by putting software on a machine is not sufficient to create an archive. In the case of an archive like E-LIS; its organisational model is the sum of its policies. An archive without policies is like a library without a librarian. As example the legal framework is a main task: policies and practice for the management of intellectual property of data and the provision of open access to data and literature are needs to support the ability of the research communities to share, access, and reuse data, as well as to integrate data from diverse sources for research, education and other purposes. High-level recommendations will help research funders, infrastructure managers, research and cultural institutions and researchers for all the disciplines in consideration in furthering the goal of open data and open access in their organization and network and establish a harmonized policy for sharing and reuse data. Privacy, sensitive and personal data are considered

by administrave staff as matter particularly delicate. We all know that from research data you might identify individuals or obtain sensitive information. That's why authors must be aware of these risks and handle all these information in a secure and law-compliant way.

All the continents are now represented, over 120, with a distribution that actively involves more than 60 different countries, where there is available an editor for such Country. In ELIS it is possible to browse by country. This gives a truly international aspect to the archive and is particularly aligned with the organisation of the editorial board whereby work is channelled through international staff on an individual country basis.

- AFRICA (150)
  - Algeria (1)
  - Angola (1)
  - Botswana (3)
  - Cameroon (2)
  - Central African Republic (1)
  - Egypt (5)
  - Ethiopia (5)
  - Ghana (3)
  - Kenya (15)
  - Lesotho (1)
  - Madagascar (1)
  - Malawi (1)
  - Morocco (14)
  - Namibia (2)
  - Nigeria (29)
  - Senegal (2)
  - Seychelles (1)
  - South Africa (62)
  - Sudan (1)
  - Swaziland (1)
  - Tanzania (5)
  - Tunisia (2)
  - Uganda (6)
  - Zambia (5)
  - Zimbabwe (11)
- OCEANIA (155)
  - Australia (112)
  - Melanesia (2)
    - Fiji (1)
    - Solomon Islands (1)
  - New Zealand (44)
- AMERICA: North and Central America (3188)
  - Antigua and Barbuda (2)
  - Canada (491)
  - Costa Rica (180)
  - Cuba (789)
  - Dominican Republic (9)
  - El Salvador (6)
  - Guatemala (6)
  - Honduras (2)
  - Jamaica (1)
  - Mexico (803)
  - Nicaragua (7)
  - Panama (1)
  - Puerto Rico (29)
  - Trinidad and Tobago (13)
  - United States (914)
- AMERICA: South America (3645)
  - Argentina (1296)
  - Bolivia (32)
  - Brazil (1008)
  - Chile (263)
  - Colombia (647)
  - Ecuador (38)
  - French Guiana (2)
  - Guyana (1)
  - Paraguay (7)
  - Peru (228)
  - Suriname (1)
  - Uruguay (90)
  - Venezuela (122)
- ANTARCTICA (2)
- ASIA (1937)
  - Azerbaijan (1)
  - Bahrain (1)
  - Bangladesh (51)
  - China, People's Republic of (176)
  - Hong Kong (2)
  - India (1090)
  - Indonesia (136)
  - Iran (205)
  - Iraq (2)
  - Israel (6)
  - Japan (9)
  - Kuwait (5)
  - Kyrgyzstan (1)
  - Lebanon (11)
  - Malaysia (88)
  - Nepal (16)
  - North Korea (1)
  - Oman (1)
  - Pakistan (81)
  - Philippines (24)
  - Saudi Arabia (14)
  - Singapore (1)
  - South Korea (3)
  - Sri Lanka (24)
  - Syria (2)
  - Taiwan (15)
  - Thailand (12)
  - United Arab Emirates (1)
  - Vietnam (3)
- EUROPE (12127)
  - Austria (944)
  - Belarus (6)
  - Belgium (98)
  - Bosnia Herzegovina (10)
  - Bulgaria (124)
  - Croatia (118)
  - Cyprus (48)
  - Czech Republic (158)
  - Denmark (19)
  - Estonia (8)
  - Finland (17)
  - France (157)
  - Germany (690)
  - Greece (530)
  - Hungary (2)
  - Ireland (11)
  - Italy (1701)
  - Latvia (3)
  - Lithuania (42)
  - Luxembourg (1)
  - Macedonia, Republic of (2)
  - Moldova (1)
  - Norway (15)
  - Poland (483)
  - Portugal (269)
  - Romania (32)
  - Russia (14)
  - Serbia (22)
  - Serbia and Montenegro (243)
  - Slovakia (2)
  - Slovenia (31)
  - Spain (5069)
  - Sweden (30)
  - Switzerland (136)
  - Turkey (509)
  - Ukraine (245)
  - United Kingdom (594)
  - the Netherlands (86)

120 Countries present, 27 different languages

Illustration 8.

The collaboration with countries ignored for years by the librarianship tradition is what has made E-LIS particularly innovative on issues previously poorly represented or considered on the margins. Topics that reflect a "different" cultural approach in E-LIS finds space, generating an intellectual growth with respect to the comparison between identity and otherness, in particular respects to the presence of Grey Literature. The studies on the bibliometrics of Indian colleagues, the collaboration with Cuban librarians at the time of the US embargo, the request to include the Maori language by New Zealand colleagues, the emergence of contents from the East of the world, the ferment of the South American jobs for Open Access are just some of the inclusiveness traits of the multicultural character of E-LIS, one of the reasons for its success. Its organizational structure of international scope makes it a model for the construction of open digital libraries, exportable to other communities.

Data from:

"E-LIS between old and new forms of Grey Literature encompasses new forms of relationship between librarians in different countries"

https://doi.org/10.17026/dans-xg3-ty57

Antonella De Robbio, e-LIS ePrints for Library and Information Science
https://orcid.org/0000-0001-6617-444X
Dataset Creator / Draft Text

Imma Subirats, e-LIS ePrints for Library and Information Science
https://orcid.org/0000-0002-3401-8597
Visualization

Fernanda Peset, e-LIS ePrints for Library and Information Science
http://orcid.org/0000-0003-3706-6532
Visualization

Abstract

E-LIS is an international digital repository for Library and Information Science (LIS), including Communication. Created in 2003 and hosted by Università di Napoli Federico II in Italy, in a few years has been indeed as the largest international open repository in the field of library and information science. The strength of our archive is that all the work is completely based on voluntary work. It has grown thanks to a team of 80 volunteer editors, LIS professionals as technicians, librarians and information specialists. The editorial team is formed by 67 editors coming by different countries plus a team of technicians and the Administrative Board which decides policies and rules. Data are divided into general categories, according to the way they have been collected or elaborated, and these are the most common types. One of the 23 typology of metadata concerns "dataset" which can be described and deposited as a item itself inside the repository.

After 16 years from its take-off, the disciplinary repository contains 22.000 open access contents in 27 different languages. The set contains statistical data about E-LIS. Over 20,000 papers have been deposited rigorously as full text open access, with a media of 110 deposits per months. Mainly are constituted by journal articles. The geographical distribution of E-LIS deposits by Countries is subdivided by countries.

The possibility to get statistics give us the dimension of the real interest by users coming by all the world. There are over a million annual download requested by users all over the world, numerous from the United States but also from China and South America. This data means that the repository is live and well knowing not only by LIS communities, also it is a referral point for research of papers published and Grey Literature.

Keywords: International Open Archive; Library and information science; Repository; Open Access; Data Paper; Policy Development; Grey Literature

Subject Area: Library Information Science; Grey Literature

Methods Applied:

• Steps

In order to extract statistical data we used the internal statistical system of e-LIS. Data documented the vitality of the international repository. We extracted four data sets:

1. Papers deposited per months since 2002 – Format .Json
2. Papers download by users in the world – Format .Json
3. Geografic distribution of deposited papers present into e_LIS – Format .PDF
4. Image of e-LIS map refers to countries of e-LIS team – format JPG

• Sampling strategy

All of the continents of the world are represented with submissions from 120 countries of which 60 have a dedicated editor. On average 110 deposited publications are validated each month.

Dataset Description

File name(s):	
1. *E-LIS_Deposit_per-Months.json*	
2. *E-LIS_Download_data.json*	
3. *E-LIS_GeograficDistributionPapersDeposited.pdf*	
4. *E-LIS_TeamWWW.JPG*	
***Format name and version*:**	.JSON; .PDF; JPG
Creation dates:	from 2002 to 2019
Language:	English
License:	CC0 Waiver - no rights reserved
Archive name:	DANS EASY Archive
Publication date:	2020-04-10

Potential Reuse of the Data

E-LIS indexes intellectual works in the field of librarianship and information science. All bibliographic data, produced by the two main participating entities in this indexing processes, namely the submitters and the editors, are open according to the Open Data Commons Open Database License. In support of this practice, e-LIS endorses the OpenBiblio Principles as published in January 17, 2011. Third parties may collect bibliographic data from e-LIS via automated mechanisms and facilitate end-user services to support the dissemination and retrieval of the archive's content. E-LIS general policy is to allow the harvesting of bibliographic data, but explicitly prohibits the automated harvesting of the full content of the intellectual works. The data collected in this survey are statistical data about e-LIS and it permits potential reuse and further analysis not only because all rights have been waived and the data is publicly accessible, but also because of its interest to other communities of practice pertaining to long-tail research on policy development. This data paper demonstrates compliance with the FAIR principles: Findable, Accessible, Interoperable and Reusable. The same applies for policies, tools, and infrastructure of e-LIS that interoperate in fulfilling these principles.

Linked References

http://eprints.rclis.org/information.html
http://eprints.rclis.org/cgi/stats/report/deposits?range=_ALL_&from=&to=
http://eprints.rclis.org/cgi/stats/report/requests?range=_ALL_&from=&to=
http://eprints.rclis.org/cgi/stats/report/compare_years?range=_ALL_&from=&to=

Grey literature as Valuable Resources in the National Library of Iran: From Organizing to Digitization*

Reza Shahrabi Farahani and Somayeh Sadat Hashemi,
National Library and Archives of Iran

Abstract
Grey literature is an important source of information due to the uniqueness of the content that gets published. Because commercial publishers are looking to make a profit on the materials they publish, they often overlook niche research areas that serve smaller populations. Grey literature is one way to search for information in emerging or less popular research areas. It seems that it reinforce research and accelerate continuous movement towards having a scientific society. Any source of information that has not been released to the market after printing is considered as grey literature. Grey literature is a special resource that is created for a specific purpose and audience and is replicated finitely. It is impermanent and invisible, and cannot be seen in the directory of publishers, bookshops and libraries.
The National Library of Iran was founded in 1937. The main goal of this organization is collecting, preserving, organizing and disseminating information about printed and non-printed works in Iran, and taking measures and making decisions to guarantee the accuracy, ease and speed of research and study in all fields to promote national culture. In order to achieve these goals, and according to the law, all private and public publishers are required to submit a copy of their publications (book and non -book materials) to the National Library of Iran . There are more than 900 thousand issues in the grey literature group at the National Library, over 240,000 issues of which are digitized. After being collected, these resources are organized and made available to users. Report, research project, standard and dissertations are among the most important sources of the National Library of Iran 's grey literature, used by many researchers daily.In this research, the process of collecting, organizing and disseminating of information of these sources in the National Library of Iran will be expressed as a successful practical experience. Besides, it will be shown how the National Library of Iran has dealt with the problems it faced with.
Keywords: Grey literatures, Collecting, Organizing, Digitizing, Dissemination of information, National Library of Iran

Introduction

Whenever concern is expressed at the continually increasing quantity of grey literature, and at the difficulty it presents to librarians, information workers and documentation experts on the one hand, and to readers and users on the other, the reply which is often given is that such publications are not intended to form part of the permanent literature, and consequently any problems that may arise will be purely temporary. Those who prepare and issue grey literature do so because such documents offer a number of advantages over other means of dissemination, including greater speed, greater flexibility and the opportunity to go into considerable detail if necessary. Another important reason why grey literature has attained significance as a separate medium of communication is because of an initial need for security or confidentiality classifications which prevent documents from being published in the conventional manner. Over the years, grey literature has come to constitute a section of publications ranking in importance with journals, books, serials and specifications. Any subjective feelings that grey literature is being used and quoted more and more are borne out by the many independent references to it, and by the emergence of databases devoted specifically to standardizing its identification and improving its accessibility.

* First presented at the 2019 IFLA WLIC Satellite Meeting: Grey Literature: Scholarly Communication in a Digital World August 23, 2019 at the National Library of Greece @ SNFCC (Stavros Niarchos Foundation Cultural Center), Athens, Greece

What is Gray Literature?

The problem of trying to define items of grey literature is one which is well known to the life scientist – it is rather like trying to define the species. Everyone can recognize a piece of grey literature when they see it, but it is not easy to write an explanation which covers all the exceptions (Auger, 1997).

The Fourth International Conference on Grey Literature (GL'99) in Washington, DC, in October 1999 defined grey literature as follows: "That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers."

In general, grey literature publications are non-conventional, fugitive, and sometimes ephemeral publications. They may include, but are not limited to the following types of materials:

- Reports (pre-prints, preliminary progress and advanced reports, technical reports, statistical reports, state-of-the art reports, market research reports, government reports, etc.)
- Theses and dissertations.
- Conference proceedings
- Technical specifications and standards
- Non-commercial translations
- Bibliographies
- Policy statements and issues papers.
- Pre-prints and post-prints of articles.
- Geological and geophysical surveys.
- Maps.
- Newsletters and bulletins.
- Fact sheets.
- Technical and commercial documentation, and official documents not published commercially (primarily government reports and documents) (Alberani, 1990).

Characteristics of Grey Literature

1. Not primarily produced for commercial publications
2. Production as a means of getting a message across rather than publishing as a commercial venture (for project or otherwise)
3. Difficult to acquire
4. Not part of a major distribution channels
5. Few, if any bibliographic controls e.g. lack of ISBN or ISSN
6. Not peer reviewed
7. Transient or ephemeral in nature
8. Difficult to find because historically, it is not included in commercial abstracting and indexing database

The importance of grey literature

Much grey literature is of high quality. Grey literature is often the best source of up-to-date research on certain topics (Grey literature, 2019).

Grey literature is an important source of information. It can often be produced more quickly as it has greater flexibility. It serves scholars and lay readers alike with research summaries, facts, statistics, and other data that offer a more comprehensive view of the topic interest.

The following are some of the importance of grey literature publications:

- ✓ It can provide information that often is unavailable in published open sources
- ✓ It is often available on timelier basis than conventional literature.
- ✓ Conference papers, for example, are available long before any follow-on, published article will appear, yet the information content of the two versions may not differ significantly.
- ✓ It can corroborate important assertions found in other sources, which is always paramount in intelligence analysis.

- ✓ It may have a concise, focused, and detailed content. This is particularly true of technical reports and unofficial government documents, whose information content will be greatly reduced in the published form.
- ✓ The literature is always free, relevant and unique.

It is becoming a common means of information exchange, particularly as personal publishing software improves and Internet access expands (Soule and Ryan, 1999).
Grey literature is an important source of information due to the uniqueness of the content that gets published. Because commercial publishers are looking to make a profit on the materials they publish, they often overlook niche research areas that serve smaller populations. Grey literature is one way to search for information in emerging or less popular research areas.
Grey literature can sometimes be more current that commercially published information. It does not go through the potentially time-consuming peer-review process undertaken by commercial publishers, and therefore has a quicker turnaround time for dissemination. Other benefits are that grey literature is more likely to report studies that ceased prematurely, as well as innovative pilot projects.

Grey literature in National Library of Iran

The National Library and Archives of Iran are a scientific, research, and service providing institution which was established in 1937. According to its legal duties, the National Library is responsible for acquiring and preserving the following resources through deposit, exchange, donation and purchase. every book and non-book material which is published in Iran, or by Iranian living abroad, all materials in the field of Iranian and Islamic studies especially about Islamic Republic of Iran, as well as scientific, technological, and cultural works published in different countries and in different languages.

- Deposit; according to the law, all private and public publishers are required to submit a copy of their publications (book and non-book materials) to the National Library of Iran.
- Exchange and Donation; some of the resources of the library are provided by exchange of materials with National Libraries ,other libraries and academic and research centers in Iran and the world. In addition, some resources are provided through donation of legal and real entities and public or private organizations.
- Purchase; purchase of books and periodicals mainly in languages other than Persian on Islamic and Iranian studies, studies related to Islamic Republic of Iran and also reference and scientific works is one of the ways for collection development in the National Library of Iran.

The main goal of this organization is collecting, preserving, organizing and disseminating information in the form of printed and non-printed works in Iran, as well as taking measures and making decisions to guarantee the accuracy, ease and speed of research and study in all fields. In order to achieve these goals, collection management is performed through various means such as cataloguing, classifying, documentation, indexing and abstracting according to international standards. This policy has resulted in fast and comprehensive retrieval of library materials.
There are more than 3 million and 500 thousand titles of information sources in the National Library of Iran, of which more than 993,000 titles (about 28 % of the total resources) are classified in the grey literature resources group. These resources include theses, research projects, standards, pamphlets, posters, leaflets, etc., which are used by a large number of researchers daily. Table 1 lists Frequency distribution of different types of grey literature resources in National Library of Iran.

Table1. Frequency distribution of different types of grey literature in NLI[1].

NO.	Types of grey resources	Frequency	Percentage
1	Announcement	14578	1.47
2	Brochure	15806	1.6
3	Bulletin	1630	0.26
4	Directory	5	0.01
5	Dissertation and thesis	306700	30.86
6	Document	553506	55.69
7	Newsletter	1235	0.21
8	Pamphlet	13180	1.33
9	Poster	20453	2.1
10	Report	26132	2.27
11	Research plan	9897	1
12	Resume	3287	0.43
13	Standard	27499	2.77
	SUM	993908	100

As the table above shows, the largest amount of gray literature in the NLI is related to documents (%55.69) and the lowest amount is related to the directories (less than %0.01).
After the documents, theses (%30.86), standards (%2.77), reports (%2.27), posters (%2.1), brochures (%1.6), announcements (%1.47), pamphlets (%1.33), research plans (%1), resumes (%0.43), bulletins (%0.26), and newsletters (%0.21) have a largest number of grey literature in the NLI.

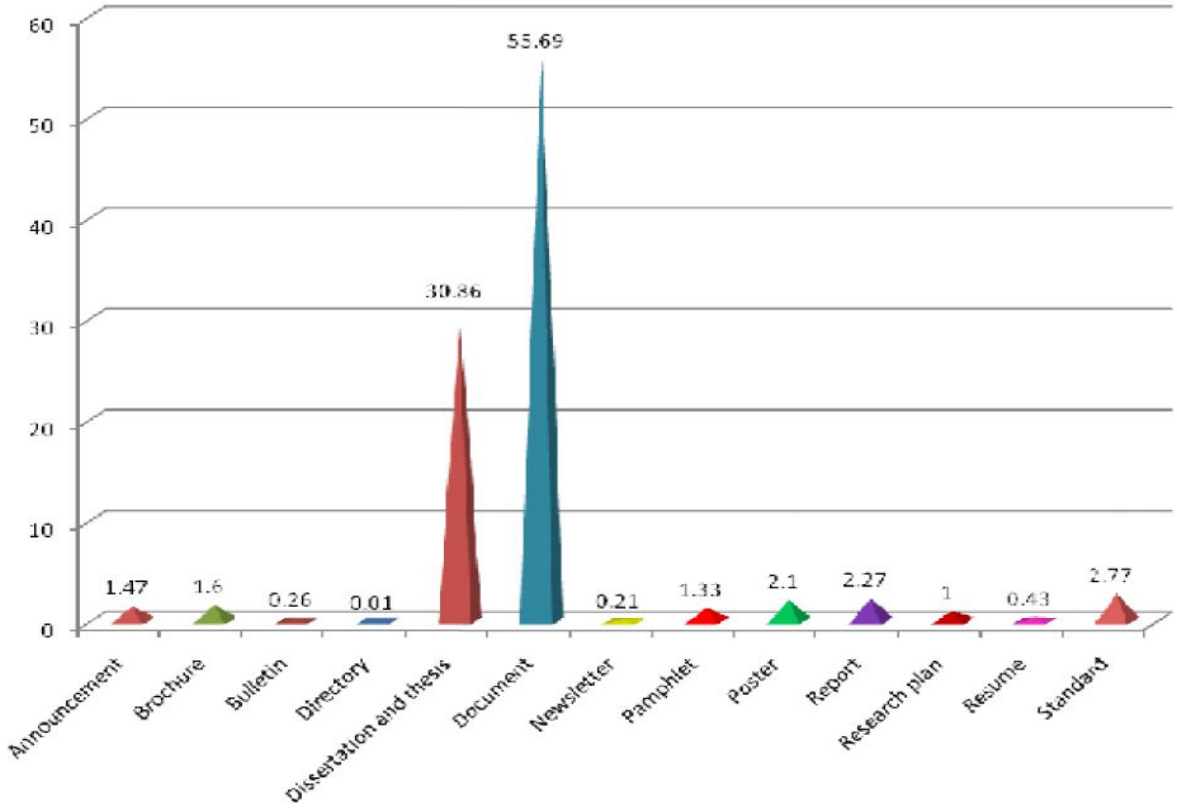


Figure1. The percentage of different types of grey literature in NLI

Grey literature are collected and organized in the National Library of Iran through the depository law, purchase, exchange and donation.The depository law of non-book resources was approved in 1999.Accordingto the law, all governmental and nongovernmental producers of non-book

[1] National Library of Iran.

resources have to deliver two copies of their works to the National Library of Iran. One of the most important of these resources is dissertations and theses that students and researchers use in the NLI. Approximately 150 graduate and postgraduate researchers come to the National Library for their academic research and use these resources daily. Another gateway is exchange and donation of resources to the NLI through which the library resources are added. After collecting grey literature, the first step to organize these resources is processing and cataloguing. This process help researchers to access the resources by searching in the library software.

Cataloguing of resources in National Library of Iran is based on the standard named IRANMARK which is based on UNI MARK and customized.

UNIMARC is a specific implementation of ISO 2709, an international standard that specifies the structure of records containing bibliographic data. It specifies that every bibliographic record prepared for exchange conforming to the standard must consist of:

A RECORD LABEL consisting of 24 characters,

A DIRECTORY consisting of a 3-digit tag of each data field, along with its length and its starting character position relative to the first data field,

DATA FIELDS of variable length, each separated by a field separator.

Resource cataloguing based on MARK does not have long history in NLI. In addition, cataloguing according to this standard and completing all information blocks for each record is an extremely time-consuming process and due to the large number of resources received daily by the NLI, most of them are catalogued descriptively and their analytical cataloguing –including indexing- is being completed gradually. Descriptor, scope and content, electronic location and access and added entries are added based on IRANMARK in the analytical cataloguing of grey literature.

The main goal of organizing resources in the NLI is dissemination of information and resource accessibility for users. To achieve this goal, the information of each source is entered in RASA software – comprehensive software of National Library of Iran –so that users can access bibliographic information inside and outside the library via OPAC[2].

In recent years, digitization project have begun to facilitate users' access to resources.

Theses and documents are one of the most important grey literatures in the NLI that have been digitized and located in digital library[3] of National Library of Iran. So far, more than 400,000 titles of grey literature have been digitized.

Table 2 presents the frequency and percentage of grey literature that has been descriptively catalogued, analytically catalogued and digitized separately.

Table2. Methods of organizing grey literature in the NLI

methods / Statistical indicators	Descriptive cataloguing	Analytical cataloguing	digitizing
frequency	993908	688593	400152
percentage	100	69.28	40.26

As the table above shows, all types of grey literature have descriptive cataloguing. This part of cataloguing includes the title and the author's name, publishing status, appearance, and notes, including abstract.69.28% of these resources also have analytical cataloguing. This part of cataloguing includes subject or descriptor, scope and content, electronic location and access as well as added entries.Only 40.26% of the resources have been digitized.

[2] http://opac.nlai.ir

[3] http://dl.nlai.ir

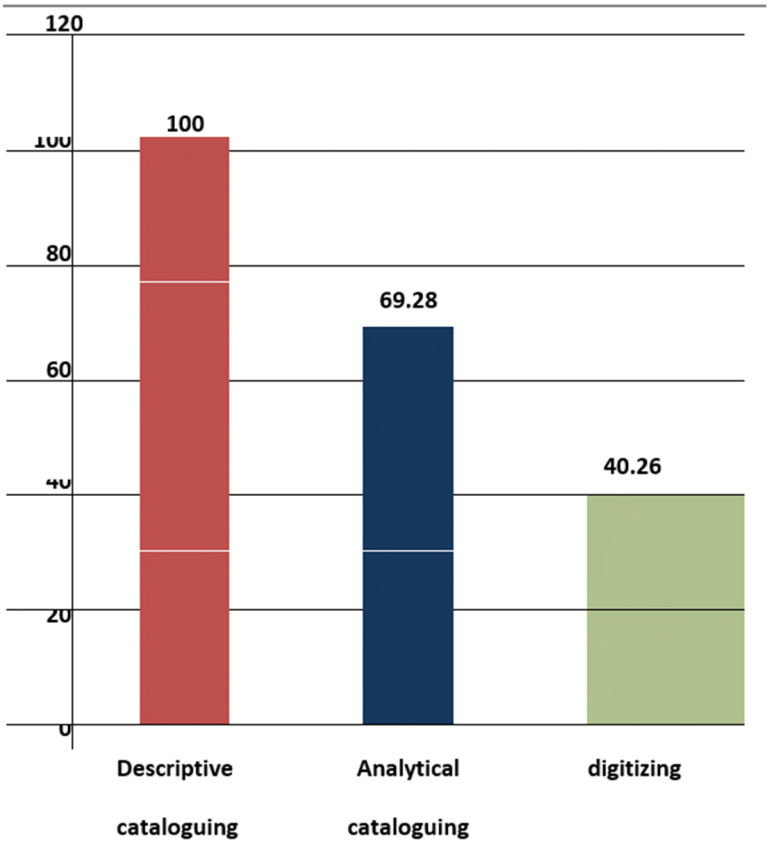


Figure2. Methods of organizing grey literature in the NLI

Users require presenting in the National Library of Iran to study grey literature. They can study these resources both in print and in digital formats. Only the bibliographic information of these resources can be searched through the OPAC, and users can only access the digital format of these resources through digital section inside the NLI (due to compliance with copyright law).Table 3 presents accessibility methods to grey literature in the NLI.

Table3. Accessibility methods to grey literature in NLI

Accessibility methods Statistical indicators	Only printed format	Only digital format	Both of them
frequency	599156	329594	70558
percentage	59.95	32.98	7.07

As Table 3 shows most access to resources is in printed format with %59.95.Only %32.98 of these resources is in digital format. Access to both printed and digital formats is %7.07.This is only for researchers who want to study the original resource.

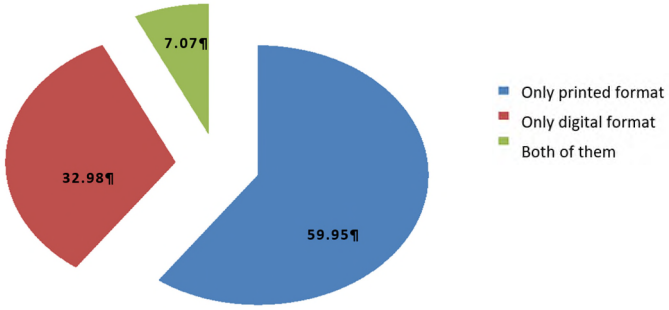


Figure 3: The percentage of different types of accessibility methods

According to the loan service survey, the amount of loaning and studying grey literature is about 20% of the total resources that were loaned by users in the NLI.175,000 grey literature loaned by users in the NLI over the past year (from April 2018 to April 2019).Most requests are related to the theses, documents and reports.

The lack of search for grey literature based on the descriptor or subject due to their incomplete cataloguing is one of the reasons for the imbalance in the use of them by users, so that users in some cases are unaware of such resources in the NLI unless the reference librarians help them to identify resources needed.

Conclusion

On the one hand to identify these resources and follow-up receiving them by the NLI is difficult because of the lack of an official publisher, the difficulty of locating, documents not being officially registered in anywhere, the lack of a standardized template, and the lack of ISBNs or ISSNs. On the other hand, these have led scholars to call them ephemeral resources.

One of the reasons for the lack of some of these resources, such as directories, is the lack of an accurate and comprehensive definition of these resources, which has led not to be catalogued in the book database.

In order to have a rich collection of grey literature, it is necessary to pay attention to all methods of collecting these resources such as depository, purchase, exchange and even direct communication with private and public publishers. Therefore, it is necessary to develop a longterm and continuous program for providing all grey literature in the National Library of Iran. It should be noted that only the dissertations of Islamic Azad University and the government documents are deposited to the National Library of Iran.

All grey literature are not organized in the same way. For example organizing theses, documents and standards are done completely, but the others are not. The information content, the lack of formal publication and the lack of formal distribution channels can be regarded as the causes.

Finally, the following proposals are suggested:

- ✓ Preparing an instruction for the collection, organization, and dissemination of grey literature;
- ✓ Establishment of a grey literature selection committee;
- ✓ Preparing the union catalogs of grey literature;
- ✓ Collecting different types of grey literature from other organizations;
- ✓ Holding organizing grey literature training courses for librarians.

References

- Alberani V, Pietrangeli PDC, Mazza AMR (1990). The use of grey literature in health sciences: a preliminary survey. Bulletin of the Medical Library Association 78(4)
- Auger, Peter (1997). Information sources in Grey Literature. London: Bowker.
- Grey Literature'99 Conference Program. Fourth International Conference on Grey Literature: New Frontiers in Grey Literature.GreyNet, Grey Literature Network Service. Washington D.C. USA, 4-5 October 1999.
- Grey Literature Typology. Retrieved in https://guides.library.unisa.edu.au/c.php?g=493934&p=3626645
- Grey literature: What is it? The value of grey literature. Retrieved in https://libguides.royalroads.ca/greylit/value
- National Library of Iran official website: www.nlai.ir
- Soule, Mason, H. and Ryan, R Paul (1999). Grey Literature: Technical Briefing Information.Information Technology summit www.dtic.mil/summit/tb07.html

Data from:

"Grey literature as Valuable Resources in National Library of Iran: From Organizing to Digitization"

https://doi.org/10.17026/dans-zhp-ncke

Somayeh Sadat Hashemi
https://orcid.org/0000-0002-7728-1600
Project Administration / Dataset Creator / Draft Text

Reza Shahrabi Farahani
https://orcid.org/0000-0002-6649-1610
Reviewer / Analyzer

Abstract

The National Library of Iran (NLI) was founded in 1937. The main goal of this organization is collecting, preserving, organizing and disseminating information about printed and non-printed works in Iran. This dataset is based on the grey literature available in the National Library of Iran until April 4, 2020. There are 1,120,999 grey literature documents in the National Library of Iran of which 524,162 are digitized. The dataset consists of six questions, the last of which is open- ended and all of which allow for additional comments. The first two questions deal with cataloging. The third question refers to the digitalization of these resources. The following three questions indicate the accessibility to grey literature. In this study, the process of collecting, organizing and disseminating information about these resources in NLI prove to be a successful practical experience. The problems that NLI encountered are also addressed.

Keywords: Grey literatures, Collection Development, Organization, Digitizing, Dissemination of information, National Library of Iran

Subject Area: Information Science

Methods Applied

● Steps

There are more than 4,500,000 titles of information resources in the National Library of Iran of which more than one million titles (about 28% of the total resources) are classified under grey literature shown in the Table below[1]. These resources include theses, research outputs, standards, pamphlets, posters, leaflets, etc. used by a large number of researchers daily. The dataset is based on resources available in the National Library of Iran until April 4, 2020. There are 1,120,999 grey literature documents in the National Library of Iran, of which 524,162 are digitized. The questionnaire contains six queries, the last of which is open ended and all of which allow for additional comments.

● Sampling strategy

To compile this dataset, a list of grey literature in the National Library of Iran was extracted, and then the questionnaire that contains six queries was designed and the survey was carried out using RASA[2] software – comprehensive NLI software. The following table lists Frequency distribution of different types of grey literature in National Library of Iran.

NO.	TYPES OF GREY RESOURCES	FREQUENCY	PERCENTAGE
1	Announcement	14609	1.3
2	Brochure	16086	1.43
3	Bulletin	1630	0.14
4	Directory	6	~0.005
5	Dissertation and thesis	311149	27.75
6	Document	671806	59.92
7	Newsletter	1235	0.14
8	Pamphlet	13302	1.18
9	Poster	20473	1.82
10	Report	27231	2.42
11	Research plan	11402	1.05
12	Resume	3310	0.29
13	Standard	28760	2.56
	SUM	1120999	100

Dataset Description

File name:	Dataset-NLI-Iran
***Format name and version*:**	.PDF
Creation dates:	2020-04-04
Language:	English
License:	CC0 Waiver - no rights reserved
Archive name:	DANS EASY Archive
Publication date:	2020-04-19

Potential Reuse of the Data

According to its legal duties, the National Library of Iran is responsible for acquiring and preserving the following resources through deposit, exchange, donation and purchase[3]. Grey literature is collected and organized in the National Library of Iran through the depository law, purchase, exchange and donation. The depository law of non-book resources was approved in 1999. According to the law, all governmental and non-governmental producers of non-book resources are required to deliver two copies of their work to the National Library of Iran. One of the most important of these resources is dissertations and theses that students and researchers use in the NLI. Approximately 150 graduate and postgraduate researchers come to the National Library for their academic research and use these resources daily. In recent years, a digitization project has begun to facilitate user access to these resources. Theses and documents are one of the most important grey literature types in the NLI that have been digitized and housed in digital library of National Library of Iran[4].

It is necessary to develop a long-term and continuous program providing all grey for literature in the NLI. Not all grey literature is organized in the same way. For example, organizing theses and standards are done comprehensively, but the others types are not. The information content, the lack of formal publication, and the lack of formal distribution channels can be regarded as the causes.

This dataset paper helps to periodically survey grey literature in the NLI. This review includes the amount of resources received, cataloging and information processing, digitalization, etc. Finally, the following proposals are suggested:

- ✓ preparing an instruction for the collection, organization, and dissemination of grey literature;
- ✓ Establishment of a grey literature selection committee;
- ✓ preparing the union catalogs of grey literature;
- ✓ collecting different types of grey literature from other organizations;
- ✓ holding organizing grey literature training courses for librarians.

Linked References

[1] . Statistics of increasing available resources in the National Library of Iran: http://www.nlai.ir/statistics-2

[2] . Bibliographic search via National Library of Iran's official website: http://opac.nlai.ir

[3] . Statute of the National Library and Archives of Iran: http://www.nlai.ir/asasnameh

[4] . Digital library of National Library of Iran official website: http://dl.nlai.ir

GreyNet

ONLINE TRAINING

Research, Publication, Open Access, Education, and Public Awareness to Grey Literature

GreyNet International is a networked organization specializing for over a quarter century in research, publication, access, education, and public awareness to Grey Literature. Grey Literature has its origins in the fields of publishing, information, and library science. Over the years, GreyNet has booked a record of achievements in organizing content related conferences, seminars, and workshops for small and medium size groups.

Now with the COVID-19 pandemic, GreyNet seeks to provide a series of online modules geared to students and faculty, as well as information practitioners and professionals. The initial course offerings include:

Four Course Modules	Training Activities
On the Heels of Open Access Publishing	**Mapping and Measuring your Capacity in Open Access Publishing**
Leveraging Information Resources	**Assessing sustained information resources among your stakeholders**
An Innovative Tool for Information Management	**Publishing your research data and writing a Data Paper**
Grey Literature and the Circular Economy: A Use Case	**Exploring the potential of grey literature in your field of study and practice**

These course offerings are tailored to the expectations of the participants and together with GreyNet form an agreed course plan. For further information, contact **info@greynet.org**

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

+31-(0) 20 331 2420
info@greynet.org
http://www.greynet.org
http://greyguide.isti.cnr.it/

Author Information

De Robbio, Antonella
Antonella De Robbio is CEO of E-LIS Eprints in Library and Information Science, the international archive and open access for librarianship and information sciences that she designed and created in 2003. He worked at the University of Padua for over 40 years, covering various roles over the years, from trainer on topics related to scientific communication, open access and rights management, coordinator of work groups on library science and e-learning and as a contact person for copyright at the University Center CAB for Libraries. After a period of work at the CERN library in Geneva in 2000, returning to Italy he was actively occupied by Open Access and in 2004 he implemented Padova @ Dissertation and Padua @ research the two institutional archives of the University for the storage of intellectual research productions and doctoral thesis. For about twenty years it has been actively engaged at national and international level in various working groups both for issues related to copyright and rights management, and in the area of Open Access. Since 2005 he is a member of the national Open Access group of the CRUI Libraries Commission. Since 2016 he is Mentor for MooC Research Data Management and Sharing (research data management) on the Coursera platform organized by the University of North Carolina and the University of Edinburgh. Within the AIB Italian Libraries Association has been for over fifteen years national coordinator of the drafting of Italian OPACs for the Italian MetaOpac MAI Azalai from 1999 to 2015 He is currently a member of the AIB Legislative Observatory and coordinator of the Public Domain Study Group and Open Access.
Email: antonella.derobbio@gmail.com
ORCID iD: http://orcid.org/0000-0001-6617-444X

Hagemann-Wilholt, Stephanie
Since August 2018, Stephanie Hagemann-Wilholt has been working for DOI Services at TIB Leibniz Information Centre for Science and Technology in Hanover. Her work focuses on advancing the use of persistent identifiers. Before that, she was a consultant at the office of the German Council for Scientific Information Infrastructures. Stephanie holds a PhD in economic history from Bielefeld University, where she was a fellow of the graduate research group at the Institute for Science and Technology Studies.
Email: stephanie.hagemann@tib.eu
ORCID iD https://orcid.org/0000-0002-0474-2410

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

Plank; Margret
Margret Plank is currently the Head of the Competence Centre for Non-Textual Materials at the German National Library of Science and Technology in Hannover (Germany). The aim of the Competence Centre for Non-Textual Materials is to develop emerging tools and services that actively support users in the scientific work process enabling non-textual material such as audiovisual media, 3D objects and research data to be published, found and made available on a permanent basis as easily as textual documents. Previously she was responsible for Information Competence and Usability at the TIB. She has also worked as a researcher at the Institute of Information Studies and Language Technology at the University of Hildesheim. She represents TIB on a number of boards including IFLA Steering Committee Audiovisual and Multimedia Section as well as ICSTI / ITOC. Margret Plank holds a Master degree in information science and media studies from the University of Hildesheim, Germany.
Email: margret.plank@tib.eu

Author Information *CONTINUED*

Savić, Dobrica
Dr. Dobrica Savić is Head of the Nuclear Information Section (NIS) of the IAEA. He holds a Doctorate degree from Middlesex University in London, an MPhil degree in Library and Information Science from Loughborough University, UK, an MA in International Relations from the University of Belgrade, Serbia, as well as a Graduate Diploma in Public Administration, Concordia University, Montreal, Canada. He has extensive experience in the management and operations of web, library, information and knowledge management, as well as records management and archives services across various United Nations agencies, including UNV, UNESCO, World Bank, ICAO, and the IAEA. His main interests are digital transformation, creativity, innovation and use of information technology in library and information services.
Email: www.linkedin.com/in/dobricasavic
ORCID iD https://orcid.org/0000-0003-1123-9693

Shahrabi Farahani, Reza
Reza Shahrabi is currently in charge of Information Services Department of the National Library of Iran. Prior to this, he served as Director of Information Technology as well as Director of Digital Resources Department. As senior digital library expert, he has gathered valuable experience in several large-scale projects. He has Implemented processes with comprehensive library management system. He has been working in the National Library and Archives of Iran (www.nlai.ir) for the past 15 years in different positions.
ORCID iD http://orcid.org/0000-0002-6649-1610

Soomai, Suzuette S.
Suzuette S. Soomai is a Postdoctoral Fellow with the Environmental Information: Use and Influence research program (EIUI) at Dalhousie University. She examined fisheries information use in decision-making in Fisheries and Oceans Canada, the Northwest Atlantic Fisheries Organization, and the Food and Agriculture Organization of the United Nations. She has been studying the enablers and barriers to communication of scientific information at the science-policy interface within and among governmental organizations. She holds a PhD and a Master of Marine Management from Dalhousie University, and an MPhil (Zoology) and a BSc (Hons) (Biology) from the University of the West Indies, Trinidad and Tobago. As a member of EIUI, Dr. Soomai completed case studies on the awareness, use, and influence of marine environmental science reports in policy-making contexts in Nova Scotia and the Gulf of Maine-Bay of Fundy region. Prior to this, she worked as a government fisheries scientist in the Caribbean.
Email: suzuette.soomai@dal.ca

Vicary, Tamsin
Tamsin Vicary works for the Aquatic Sciences and Fisheries Abstracts (ASFA) Secretariat, based at the Food and Agriculture Organization of the United Nations. Tamsin is responsible for ASFA's transition to a new business model with particular focus on increasing access to the database and improving coverage. To date this has involved the evaluation and selection of new technologies; implementing a strategy to improve coverage of grey literature; evaluation and promotion of the ASFA database and partnership. She has a Masters in Library and Information studies from University College London. Prior to joining ASFA, Tamsin was Collections Manager at the Freshwater Biological Association, and also has experience in the private sector having worked for an international law firm and professional services firm.
Email: tamsin.vicary@fao.org

Wells, Peter G.
Peter G. Wells is an adjunct professor at Dalhousie University and a senior research fellow at International Ocean Institute, Halifax, Nova Scotia, Canada. He holds a PhD in zoology from the University of Guelph. He worked as a marine scientist and aquatic toxicologist for the Canadian Federal Government for 34 years, retiring from Environment Canada in June 2006. His academic interests include marine ecotoxicology, marine information management, and community service on issues affecting the Bay of Fundy, Gulf of Maine and North-West Atlantic. He has written, contributed to, or edited over 300 primary and technical publications on water pollution and marine environmental science, including several books. He is the current editor of the Proceedings of the Nova Scotian Institute of Science. He has served on various national and international technical committees, including the National Academy of Sciences (USA), the Organisation for Economic Co-operation and Development, and the United Nations. He was a long-serving member of the United Nations Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection. He also serves as co-lead of the Environmental Information: Use and Influence research program, which has received several honors and awards.
Email: ocean2@ns.sympatico.ca

PsycEXTRA®

EXPAND BEHAVIORAL SCIENCE RESEARCH WITH THE PREMIER RESOURCE FOR GRAY LITERATURE

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




Explore original, cutting-edge, high-quality research

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

Available via EBSCO

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; 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.

Contents

'Grey Literature – A Platform and Interface for Open Science'