

# Twenty-Fourth International Conference on Grey Literature

Publishing Grey Literature in the Digital Century

## Conference Proceedings

ISSN 1386-2316



### Program Sponsors and GreyNet Associate Members



GL2022 Program and Conference Bureau

**TextRelease**

Javastraat 194-HS, 1095 CP Amsterdam, Netherlands  
<https://textrelease.com> • [conference@textrelease.com](mailto:conference@textrelease.com)  
Tel. +31 (0)20-331.2420



## CIP

### GL2022 Proceedings

Twenty-Fourth International Conference on Grey Literature “Publishing Grey Literature in the Digital Century”. - December 5, 2022 / compiled by D. Farace and J. Frantzen ; GreyNet International, Grey Literature Network Service. – Amsterdam : TextRelease, February 2023. – 148 p. – Author Index. – (GL Conference Series, ISSN 1386-2316 ; No. 24).

NLM-NIH (USA), TIB (DE), CVTISR (SK), EBSCO (USA), ISTI-CNR (IT), KISTI (KR), NIS-IAEA (UN), NTK (CZ), Figshare (UK), and the University of Florida (USA) are Corporate Authors and Associate Members of GreyNet International. These proceedings contain 13 full-text papers presented during the conference sessions, including a link to their audio-visual recordings. Also included is an author index with the names of contributing authors and researchers along with their biographical notes, ORCiDs, and ROR ID affiliations. Further contained is a list of over 30 participating organizations from 13 countries worldwide along with 15 sponsored advertisements.

# *Foreword*

## **PUBLISHING GREY LITERATURE IN THE DIGITAL CENTURY**



For over three decades, authors and researchers in the field of information have addressed the many challenges in publishing grey literature. In so doing, they have confronted core issues. Is grey literature published literature, and if so, how is it published; how does it differ from commercially published literature; and foremost how can it stay abreast with the technological developments that will ensure its access, uses, and preservation for scholarly research and citizen science well into the 21st Century.

The response to such questions lies in the collaboration and integrated roles of publishing bodies and their affiliate libraries and information centers, where grey literature is produced and processed. While currentness is inherent to grey literature in that it is situated at the cutting edge of research, and while it is comprehensive in that it captures the corpus of both the research process and issuing results, grey literature publishing often lacks the financial resources and technical expertise afforded commercial publishers.

Due consideration should be given to shared workflows grounded in an understanding and commitment that production and publication constitute two integral parts in publishing today's digital grey literature. GL2022 addresses the components of such shared workflows embedded in FAIR data principles and implemented by diverse communities of practice in this our digital Century.

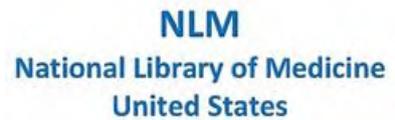
*Dominic Farace*  
GREYNET INTERNATIONAL

*Amsterdam, Netherlands*  
FEBRUARY 2023

## ***GL2022 Program Committee***



**Kristina Womack** <sup>Chair</sup>  
National Library of  
Medicine, National Institutes  
of Health  
United States



**Plato L. Smith**  
University of Florida;  
George A. Smathers  
Libraries,  
United States



**Giovanni De Simone**  
Italian National Research  
Council; Central Library  
Italy



**Brian Paul Bales**  
Nuclear Information Section,  
International Atomic Energy  
Agency, United Nations



**Margret Plank**  
German National Library  
of Science and Technology,  
Germany



## GL2022 Program Committee *CONTINUED*



**SeokJong Lim**  
Korea Institute of Science  
and Technology Information  
Korea



**Tomas A. Lipinski**  
University of Wisconsin --  
Milwaukee  
United States



**Hana Vyčítalová**  
National Library of  
Technology  
Czech Republic



**Joachim Schöpfel**  
University of Lille  
France



**Marta Dušková**  
Slovak Centre of Scientific  
and Technical Information  
Slovakia



**Dominic Farace**  
GreyNet International  
Netherlands



## ***GL2022 Program and Conference Sponsors***

**NLM**  
**National Library of Medicine**  
**United States**

**NLM, USA**  
National Library of Medicine;  
National Institutes of Health



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



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



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



**EBSCO, USA**



**NIS-IAEA, United Nations**  
Nuclear Information Section;  
International Atomic Energy Agency



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



**UF, USA**  
George A. Smathers Libraries  
University of Florida



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



**Figshare, United Kingdom**  
*Store, Share, Discover Research*

# Table of Contents

	Foreword.....	3
	Program Committee .....	4
	Program and Conference Sponsors.....	6
	Conference Program.....	9s
<b>Program</b>	I. Digital Publishing and Grey Literature.....	11
	II. Stakeholders and Policies Influencing Grey Literature.....	33
	III. Innovating and Repurposing Grey Literature.....	83
	IV. Virtual Library Instruction.....	130
<b>Advertorials</b>	EBSCO Library, Information Science & Technology Abstracts with Full Text (LISTA).....	8
	INIS, The International Nuclear Information System.....	10
	PsycEXTRA via EBSCO .....	24
	WorldWideScience.org – An International Partnership Supporting Open Science .....	76
	TIB, German National Library of Science and Technology.....	82
	NTK, National Library of Technology, Czech Republic.....	98
	CVTISR, Slovak Centre of Scientific and Technical Information.....	106
	KISTI, Korea Institute of Science and Technology Information.....	116
	Figshare - Leading Cloud Repository Software for Your Organization’s Grey Literature .....	121
	ISTI-CNR, Institute of Information Science and Technologies.....	128
	InfraScience - Infrastructures for Science, Connecting R&D to Support Open Science .....	129
	GreyGuide Portal and Repository - GreyNet’s Flagship Project 2013-2022.....	139
	GreyNet International Membership .....	140
	GL25 Conference Announcement .....	142
	GL25 Call for Papers .....	143
	PUBGREY – Publishers in Grey Literature.....	146
	TGJ - The Grey Journal, International Journal on Grey Literature .....	148
<b>Appendices</b>	List of Participating Organizations.....	141
	Author information.....	144
	Index to Authors.....	147

# *Library, Information Science & Technology Abstracts™ with Full Text*

*Available via EBSCOhost®*

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

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

#### **Comprehensive content includes:**

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

#### **Subject coverage includes:**

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

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

**Phone: 800.653.2726**

**Email: [request@ebSCOhost.com](mailto:request@ebSCOhost.com)**

**[www.ebSCOhost.com](http://www.ebSCOhost.com)**





# Conference Program

## I. Digital Publishing and Grey Literature

- Characteristics of a Well-Developed Grey Literature Repository: The Case of the International Nuclear Information System** 11  
Brian Bales, International Atomic Energy Agency; Nuclear Information Section, Austria
- A Retrospective on the Challenges of Incorporating Grey Literature into a Scholarly Publishing Platform** 17  
Alistair Reece, GeoScienceWorld, United States
- Information Fatigue Syndrome and Digital Burnout** 25  
Dobrica Savić, IAEA Information and Knowledge Management Consultant, Austria

## II. Stakeholders and Policies influencing Grey Literature

- Open Science as an Opportunity for Academic Grey Literature – A Systematic Review** 33  
Joachim Schöpfel, University of Lille and Hélène Prost, CNRS - GERiiCO, France
- Data Science as a Research Support Service and the Role of the Libraries: UF's CNI Executive Roundtable Experience** 49  
Plato Smith, Erik Deumens, and Christopher Barnes, University of Florida, United States
- Legal Implications of the CASE Act on Grey Libraries and Grey Literature Authors** 61  
Tomas A. Lipinski and Laura Christine Schein, School of Information Studies; University of Wisconsin—Milwaukee, United States
- Using the Overton policy to academic citation network: how does the policy grey literature and scholarly record connect?** 77  
Euan Adie, Terrence Bucknell, Jennifer Glover, and Ángel Luis Jaso-Tamame, Open Policy Ltd., UK

## III. Innovating and Repurposing Grey Literature

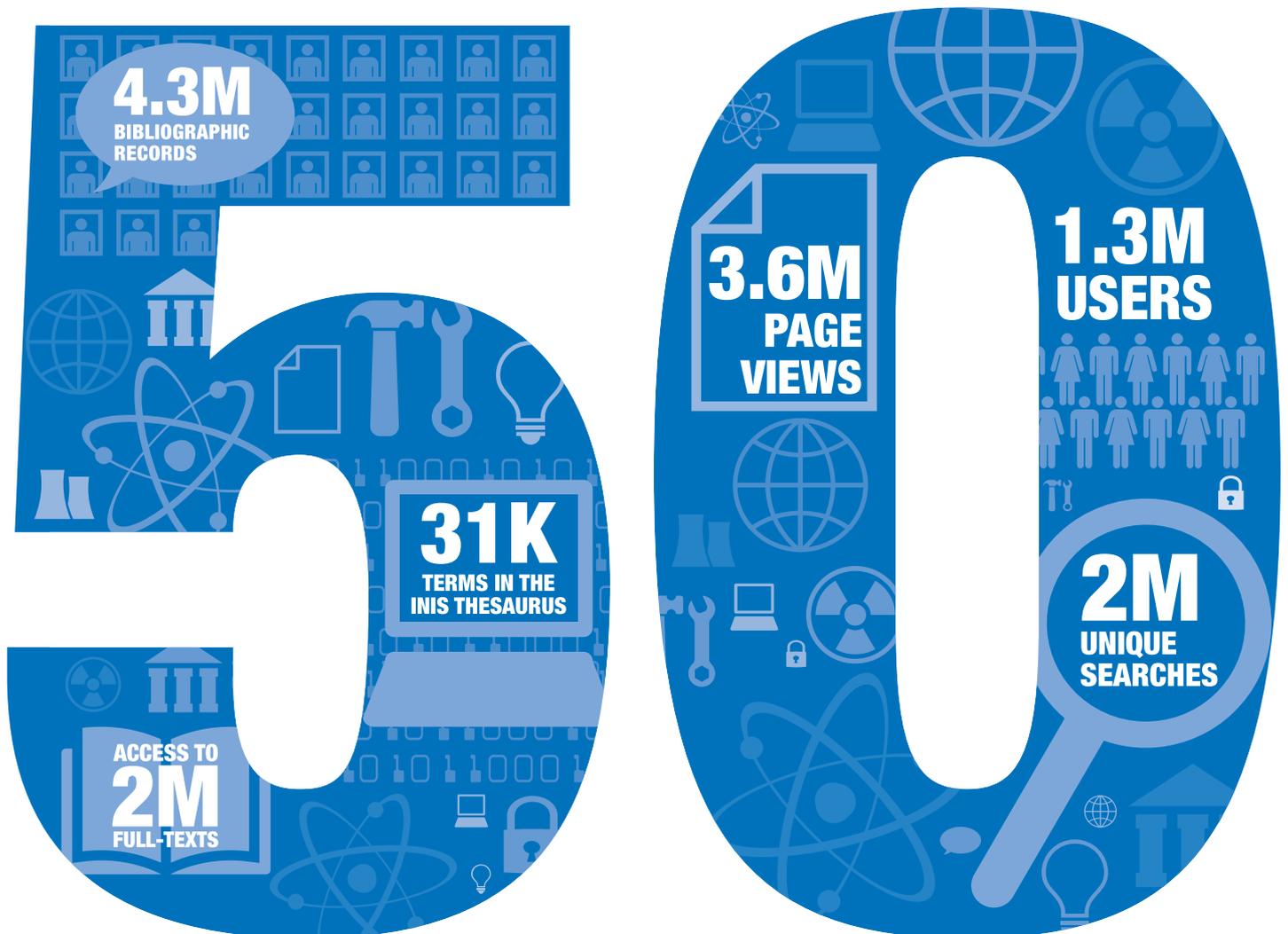
- ICSTI Member Survey: Infrastructures and Services for non-textual materials** 83  
Margret Plank, TIB Hannover, Germany; Shelby Stooksbury, OSTI-DOE, Lisa Curtin, International Council for Scientific and Technical Information, United States
- Preregistration of research for theses - a new standard?** 99  
Tereza Šímová, Czech University of Life Sciences Prague, Czech Republic
- Patent files: case study of digitalization in The National and University Library of Slovenia** 107  
Mojca Trtnik and Veronika Potočnik, National and University Library, Slovenia
- Zine Making as Autoethnographic Serious Leisure** 117  
Andrea Marshall, Centre for Media and Celebrity Studies, United States
- Digital Publishing, Open Access, and Grey Literature: The War in Ukraine 2022 as a Use Case** 122  
Dominic Farace, GreyNet International, Netherlands; Plato L. Smith, University of Florida; George A. Smathers Libraries, United States; Stefania Biagioni and Carlo Carlesi, InfraScience; ISTI-CNR, Italy

## IV. Virtual Library Instruction

- Grey is the new black: changing library instruction virtually** 130  
Aleksandra Blake and Margaret McLeod; Carleton University, Canada

# 50 YEARS OF INIS

THE WORLD'S TRUSTED NUCLEAR REPOSITORY



*Looking for nuclear information?*

*Want to preserve your nuclear information?*

**INIS CAN HELP!**

The International Nuclear Information System (INIS) was established in 1970 “to foster the exchange of scientific and technical information on peaceful uses of atomic energy”.

132 countries and 14 international organizations contribute their national nuclear literature, making it the world’s leading open access repository for nuclear science and technology literature.

**Explore** INIS and find a wealth of information on physics, radiation, climate change, health, etc. **Preserve** your nuclear information by storing it in our trusted repository.



IAEA

International  
Nuclear  
Information  
System

INIS

[INIS.Feedback@iaea.org](mailto:INIS.Feedback@iaea.org)

[iaea.org/inis](http://iaea.org/inis)

# Characteristics of a Well-Developed Grey Literature Repository: The Case of the International Nuclear Information System

Brian Bales, Nuclear Information Section,  
International Atomic Energy Agency  
<https://orcid.org/0000-0003-1452-2785>

## Video Presentation

<https://av.tib.eu/media/59873>

### Abstract

*The number of national, institutional, and subject repositories of grey literature has increased dramatically over recent years. The Directory of Open Access Repositories (OpenDOAR) currently lists 5848 repositories, a 75% increase over the last five years. Most of these repositories hold grey literature of one type or another. The degree of development of these repositories is mixed, some are of questionable quality while others are exemplars, so it is useful to define what constitutes a well-developed repository. A well-developed repository can be seen as one that meets the needs of end users, as well as the interests of authors and sponsoring organizations.*

*Characteristics, such as timeliness, openness, user-friendliness, accuracy, and completeness, are proposed as those which meet user and institutional needs and define the degree of development for a given repository. Timeliness refers to the speed at which materials in the scope of the repository are made available to the public. Openness is the degree to which material is accessible as well as shareable. User-friendliness is a subjective quality but is defined by the ease of use of the repository's user interface. Accuracy can be measured in many ways – the verisimilitude of metadata, the suitability of indexes and search results, the percentage of dead links to external resources such as full text, and other measures. Finally, completeness describes how well a repository encompasses its scope.*

*The International Nuclear Information System (INIS) has been in operation since 1970 as a repository for grey and traditional literature in all areas of nuclear science and technology. It existed before the wide adoption of information management principles, and invented methods and workflows to fulfil its mission.*

*Therefore, there are gaps between the ideal repository, embodied in the outlined characteristics, and the repository as it currently stands. These gaps are identified and solutions, as well as a plan for implementation, are proposed.*

**Keywords:** *Repositories; Grey Literature Resources; System Design; Change Management*  
**Subject Area:** *Information Science; Grey Literature.*

### Current Situation

The number and scope of Grey Literature Repositories, including national, institutional, and subject repositories, has expanded greatly in recent years. The Directory of Open-Access Repositories lists 5999 repositories as of December 2022<sup>1</sup>. This number is nearly triple that of ten years ago, when 2122 were listed. With so many repositories available for not only data collection, but also study of their design and architecture, the characteristics of a well-developed repository can be identified and defined.

A well-developed repository may be defined as one that best meets the interests of its sponsoring organization. By extension and considering that a sponsoring organization is

likely to want its users to also find the product useful, the interests and needs of end users should also be considered.

A subject repository, such as the International Nuclear Information System (INIS)<sup>2</sup>, seeks to cover the applicable publications within the defined scope of the repository. In the case of INIS, the scope is all areas of nuclear science and technology. This includes nuclear energy, but also subjects such as nuclear medicine, soil remediation, plant mutation breeding, preservation of cultural heritage items through nuclear technology, and many others.

INIS receives grey literature full-texts from its member countries and organizations. It also harvests applicable records from publishers and other repositories. The availability, number and ease of use of these resources has increased in recent years. This brings great opportunities for subject repositories, as well as caveats to be aware of.

In the 50+ year history of INIS, it has changed technology several times. It began publication in 1970 in microfiche and printed form. It has since evolved to have completely online availability with the INIS Repository Search website<sup>3</sup>. However, the capacity of the INIS workflow has remained static for several years – even in the face of an increasing number of available records.

In 1977, Hans Groenewegen and Ivan Zheludev wrote, “it is estimated that during 1977 the store of information relating to the peaceful uses of nuclear energy published around the world increase by some 80 000 documents.”<sup>4</sup> In that year, INIS added 68 831 records, or around 86% of the estimated possible input. Now, I estimate that the possible available number of records to be around 250 000 created each year in the current scope of INIS. The system is currently only capable of ingesting 125 000 per year. This workflow had a capacity in excess of the possible corpus in 1977, but in 2023, it is about half.

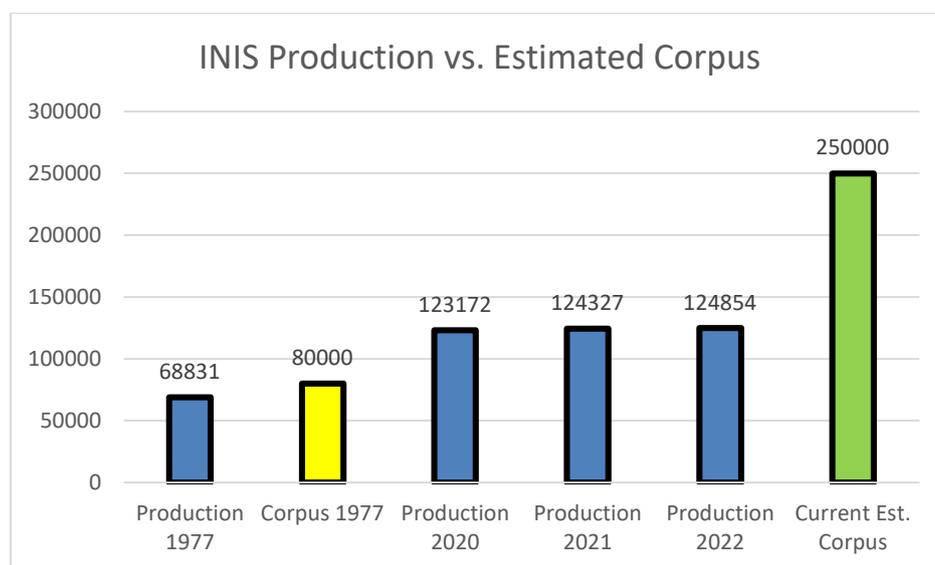


Figure 1: INIS 1977 Production vs. Estimated 1977 Corpus, Current Production vs Estimated Current Corpus

Therefore, in a time when INIS is considering redesign of its services it is good to consider what makes a well-developed repository. What exemplars exist that INIS may wish to emulate, and what can be done to better meet the needs of the sponsoring organization (the IAEA) and end users around the world.

## Exemplars

### *The Astrophysics Data System*<sup>5</sup>

Although several repositories and standards were considered for deriving the best characteristics, for the purposes of brevity, two were chosen for this paper. First, the Astrophysics Data System (ADS) is considered. Founded in 1991, the system provides over 13.3 million references. It harvests a selected number of journals with a daily frequency. That is, when an article is published in a journal covered by ADS, the reference will appear in the repository that day or the next.

Records housed in ADS always lead to a full text, though full text records are not directly available from the repository. The repository is less concerned with accuracy but more interested in speed. For quality control, ADS invites user comment and correction. ADS has an API for retrieval and allows harvesting “for personal use”, though its terms and conditions prohibit systematic downloads of even bibliographic data.

The repository has embarked upon special projects, such as harvesting core journals of not only recent articles, but also those going back to the beginnings of their publication.<sup>6</sup>

### *INSPIRE-HEP*<sup>7</sup>

This repository is the closest thematically to INIS. The scope of this repository is on high-energy physics. Although it is hosted by CERN (the European Organization for Nuclear Research), it is a joint product of several different research organizations. In operation for nearly 50 years, it houses over 1.5 million records.

Like ADS, INSPIRE-HEP harvests from core journals and repositories automatically with a daily frequency. It is less concerned with accuracy and more interested in speed of input. It invites user correction for quality control. It also has the capacity for user submission of records. It has an API for batch and automated operations. INSPIRE-HEP and CERN generally is a leader in open access. Basic search, and very specific advanced search are also available.<sup>8</sup>

## Standards

Several standards exist for scientific repositories. These include those promulgated by FAIRSFair, OpenAIRE, cOAlition S, the Core Trust Seal, and others.

The FAIRSFair project “Fostering FAIR Data Practices in Europe”<sup>9</sup>, seeks to establish its principles within what it calls the “infrastructure of science.” These principles are that scientific information and data should be findable, accessible, interoperable, and reusable. Much of this speaks to the need for openness. Openness removes financial barriers to science, where many publishers had been and continue to charge high access fees.<sup>10</sup>

The CoreTrustSeal<sup>11</sup> Trustworthy Data Repository requirements<sup>12</sup> provide 16 areas that a repository should consider and define to be able to receive the certification. The areas include preservation, security, data reuse, licensing, etc.. CoreTrust and FAIRSFair have collaborated to combine their standards and principles into a capability maturity model<sup>13</sup>.

OpenAIRE<sup>14</sup> is a coalition of several universities, to encourage open science in Europe. It has published a set of guidelines for repositories which specify how metadata should be structured for interoperability and openness. The guidelines specify the use of Dublin Core and DataCite metadata schemes and their mapping to OpenAIRE required fields. This provides a standard method for openness.<sup>15</sup>

## Characteristics

From studying the needs of INIS, as well as the configurations of exemplar repositories and current repository standards and best practices, some characteristics of well-developed grey literature repositories can be derived. The combination of this study derives the following: timeliness, openness, preservation, user-friendliness, and comprehensiveness. Combined, these form the mnemonic TOPUC.

### *Timeliness*

Timeliness is defined as “being done at a favorable or useful time.” For both traditional and grey scientific literature, the most beneficial time would be closest to the time the ideas were formulated, the discoveries made, and the paper written. The longer a publication goes without it being read, the more obsolete and therefore less valuable it becomes. Therefore, timeliness is a very important quality in a repository. Both exemplar repositories, as well as many others, are automated retrieval, processing, and publishing so records appear very quickly. This adds value for the user, who gets to access the latest information. It furthers the purpose of the organization and repository, by making it a more valuable resource.

### *Openness*

When considering setting up a repository, a sponsoring organization is seeking to collect and have a collection of literature (grey and/or traditional) be discovered and read. While some repositories are collected for internal use and analysis, INIS and other, similar, repositories were founded for the express purpose of preservation and dissemination of information. Therefore, it better meets the purpose of these repositories and the founding organization to have their contents spread as widely as possible. It also benefits the repository to have an open ecosystem, where the repository can also take advantage of the openness of still other repositories. Therefore, openness is encouraged by certifying organizations and sponsoring organizations. The ultimate benefactor is the end user community which has access to a wide variety of open information in multiple forms and multiple sources, not hidden behind a financial wall.

### *Preservation*

Like timeliness, preservation ensures that records are available now and into the future, extending to future generations. Throughout history are stories where important information was lost, and so needless suffering had to continue – sometimes for hundreds of years. The formula for roman concrete, the fact that fruit and vegetables containing vitamin C cured scurvy, and the formula for Damascus steel were all lost, ignored, or forgotten. Managers of repositories are custodians for a time to the information contained within. They cannot know which, if any, of the pieces of information they are tending will be valuable in the future. Therefore, the safest course of action would be to preserve everything possible. There is an extensive body of thought on digital preservation, along with accompanying standards, models, and capability maturity models. In brief, these account for issues such as file and format obsolescence, media degradation, malicious and accidental deletion, natural and manmade disasters, and file integrity.

### *User Friendliness*

User friendliness is the degree to which users find an interface compelling and easy-to-use. A user-friendly repository means that users will return to the interface and make it their preferred site for research. Google was able to win the search engine war by providing a simple and flexible interface that hid a great deal of complexity. One caveat is that an interface should be user friendly both to novices and expert researchers. Experts have a need to shape search queries so that specific records are returned. Amateurs need

a clear and simple interface that provides the answer they are looking for. In a modern repository, the needs of both audiences should and can be accounted for.

### *Comprehensiveness*

In essence, each of these is about extending availability. Timeliness and preservation extend the availability of records temporally. Openness extends availability physically and financially, by making records available in multiple locations and without cost. User friendliness makes records available to both novice and expert users with a compelling interface. Comprehensiveness extends the availability of records to all possible candidate records within the scope of the repository.

### **Future Work**

The list of characteristics does not purport to be comprehensive. Further characteristics could be found and added to the list. The characteristics could also be seen as a starting point, both in developing standards, and in the development of a capability maturity model.

INIS and other repositories that have not adopted these characteristics should do so. For INIS to fully adopt the characteristics of timeliness and comprehensiveness, great changes would need to be made to its workflow. Many of the processes are manually completed or at least manually executed. Furthermore, although INIS does have some capability in preservation, it does not completely meet the applicable standards. Additionally, INIS does not provide an API for batch and automated operations, and so is not interoperable with other repositories. In the coming years, all these issues should be addressed, and the characteristics adopted.

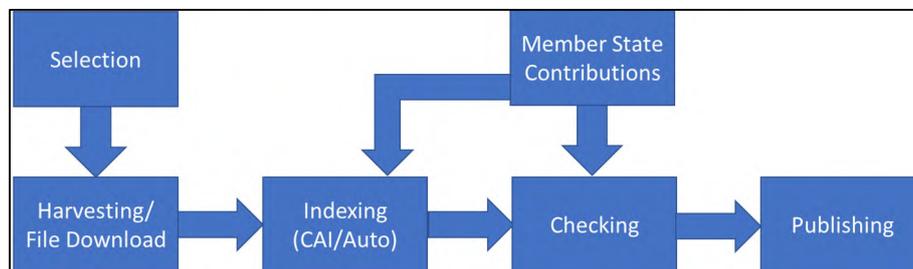


Figure 2: Simplified Current Workflow, Most Steps Require Manual Initiation, Capable of 125 000 to 130 000 per year

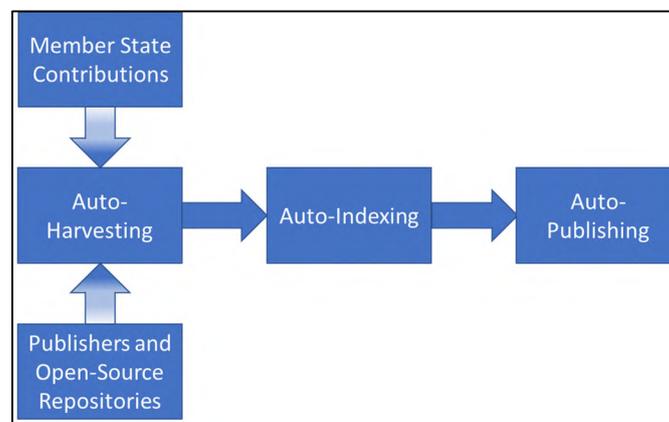


Figure 3: Proposed Future Workflow, All Steps Automated

## Conclusion

Although INIS can be considered a successfully repository, now it is over 50 years old. It is time to consider changing the system to suit modern best practices. Grey literature and hybrid repositories, including INIS, should consider adopting these 5 characteristics and emulate the model the exemplar repositories provide. In doing so, INIS will be able to fulfill the purpose for which it was created and be an indispensable repository for research in nuclear science and technology.

## References

- <sup>1</sup> Directory of Open Access Repositories, "Statistics", [https://v2.sherpa.ac.uk/view/repository\\_visualisations/1.html](https://v2.sherpa.ac.uk/view/repository_visualisations/1.html)
- <sup>2</sup> International Nuclear Information System (INIS), "Home", <https://www.iaea.org/resources/databases/inis>
- <sup>3</sup> INIS Repository Search, <https://inis.iaea.org/search/>
- <sup>4</sup> Ivan S. Zheludev and Hans W. Groenewegen, "INIS: The International Nuclear Information System", IAEA Bulletin, Vol. 20, No. 4, <https://www.iaea.org/sites/default/files/publications/magazines/bulletin/bull20-4/20405060717.pdf>
- <sup>5</sup> The Astrophysics Data System, <https://ui.adsabs.harvard.edu/>
- <sup>6</sup> The Astrophysics Data System, "About", <https://ui.adsabs.harvard.edu/about/>
- <sup>7</sup> INSPIRE-HEP, <https://inspirehep.net/>
- <sup>8</sup> INSPIRE-HEP, "About", <https://help.inspirehep.net/knowledge-base/about-inspire/>
- <sup>9</sup> FAIRSFAR, <https://www.fairsfair.eu/>
- <sup>10</sup> European Commission, "Turning FAIR into reality", <https://op.europa.eu/s/xoR2>
- <sup>11</sup> CoreTrustSeal, <https://www.coretrustseal.org/>
- <sup>12</sup> CoreTrustSeal, "Trustworthy Data Repositories Requirements", <https://www.coretrustseal.org/why-certification/requirements/>
- <sup>13</sup> FAIRSFAR, "CoreTrustSeal + FAIRenabling Capability Maturity Model", <https://www.fairsfair.eu/node/797/pdf>
- <sup>14</sup> OpenAIRE, <https://www.openaire.eu/>
- <sup>15</sup> OpenAIRE, "Guidelines", <https://guidelines.openaire.eu/en/latest/>

## A Retrospective on The Challenges of Incorporating Grey Literature into a Scholarly Publishing Platform

Alistair Reece, GeoScienceWorld, United States

### Video Presentation

<https://av.tib.eu/media/59862>

### Abstract

*In 2019, GeoScienceWorld was actively planning to bring a large content and data repository, that includes a significant proportion of highly valued Grey Literature, into our existing collection of 50+ peer-reviewed journals and over 2300 books in the geosciences. Due to various external situations, including the impacts of the COVID-19 pandemic, and an absence of community-accepted standards for Grey Literature publishing, this project has stalled.*

*GeoScienceWorld continues to investigate opportunities to bring original datasets, as well as other collections of Grey Literature, predominantly in the form of partner societies' conference proceedings and related conference materials, into our traditional research platform. We are also in the early stages of planning for a new research tool that will be truly content agnostic in bringing research and valuable insights to our primary end-user stakeholders, researchers, whether in academia or industry.*

*As an organization, GeoScienceWorld is further implementing an Agile mindset and development philosophy to bring increasingly useful, and timely, resources to our stakeholder groups. A key ceremony of all truly Agile development processes is the Retrospective.*

*In this paper, I review the initial aims of the project to incorporate a large grey dataset into our traditional scholarly literature platform and provide reflections on how both GeoScienceWorld and the wider Grey Literature community can move forward to bring such valuable datasets to audiences that both want and need, such content to advance their research.*

*For each element of the initial project, I ask the following Agile Retrospective questions:*

- *What did we do well?*
- *What could we have done better?*
- *What have we learned?*
- *What are we still puzzled by?*

*As a result of applying these questions to the initial project, I will present recommendations that both inform GeoScienceWorld's future integration and presentation of Grey Literature, as well as offer a clearer path toward greater Grey Literature acceptance within traditional scholarly platforms such as ours.*

**Keywords:** *GeoSciences, publishing platform, XML, search, business models, Agile transformation*

This paper is licensed under the Creative Commons license: CC-BY-ND-3.0 (<http://creativecommons.org/licenses/by-nd/3.0/de/deed.en>).

### Introduction

In 2019, GeoScienceWorld was working on a project to incorporate a large body of content into our research platform. The content set, whilst largely traditional, peer-reviewed, scholarly journal articles and books also included a substantial amount of Grey Literature, around 30% in total. Of this subset of the content, the majority consisted of

conference proceedings and meeting abstracts, though there was also more non-textual content like maps and posters.

At the 12th Conference on Grey Literature and Repositories in Prague, Czechia, I gave a presentation titled "The Challenges of Incorporating Grey Literature into a Scholarly Platform", which gave a broad overview of the project. In the subsequent three years, this project has stalled, largely due to the COVID-19 pandemic and its impacts on the academic publishing world.

GeoScienceWorld is undergoing an Agile Transformation of our organization, primarily focusing on making the 4 Values of the Agile Manifesto the guiding principles of our organizational culture. Having practiced Agile methodologies at previous workplaces, I decided it would be a valuable exercise to have a retrospective on the project to bring Grey Literature to our platform.

### **Background**

Founded in 2004, GeoScienceWorld hosts more than 50 journals, as well as more than 2300 ebooks, from many of the leading learned societies in the geosciences. Our founding mission was to bring together peer-reviewed, society-led, research on an online platform that would encourage collaboration among the societies to benefit the whole collective, thus giving smaller societies access to the benefits of online publishing, whilst maintaining their independence.

Conversations to bring the large set of Grey Literature into the platform began in 2018, with a project plan being in place by the autumn of 2019. The onset of the COVID pandemic in early 2020 put the project on hold, where it remains, intending to restart in 2023 if conditions are right. As part of our tentative planning for the possibility of the project restarting, it was decided that we perform a retrospective on the elements of the project as laid out in the presentation given in Prague, namely:

#### **Content Preparation**

- Implications for Search
- New Business Models

In the retrospective we asked the following questions:

- what went well?
- what could be better?
- what have we learned?
- what questions remain?

#### **Content Preparation**

##### *What went well?*

Given the nature of the project, which would have seen the overall amount of content on the GeoScienceWorld platform almost double, we needed to have a thorough understanding of the scope of content to be ingested.

Working with Silverchair, our platform partner, we created a detailed breakdown of the content, its various types, formats, as well as a count of individual content pieces. This content manifest allowed us to fully grasp not just the scope of content, but also the scope of basic development that would be required to fully onboard the content within our platform's site structure. It was in the creation of this manifest that the proportion of Grey Literature became clear.

Having a reliable content manifest that included content types we knew that there would be development work necessary to support the loading and ingestion of the Grey Literature elements of the content set. Identifying the scope and types of Grey Literature within the set allowed us to avoid many of the unexpected surprises of a content migration. As such, GeoScienceWorld has implemented a thorough content discovery process into our acquisition workflow so that we have a proper understanding of the

content being brought into our aggregation. This process is being applied to all new content acquisitions regardless of the content type involved, and has allowed us to better plan project timelines and set expectations with new publishers.

#### *What could be better?*

There was one area in particular with regards to the content preparation process that we really could have done much better, and that was having an awareness of the absence of a common Document Type Definition (DTD) for creating XML files to load Grey Literature. Within traditional scholarly publishing there exist standard DTDs for both journal articles and ebooks, JATS and BITS respectively. These DTDs are used by all the major publishing platforms for loading content, making a certain level of interoperability available during migrations from one platform to another. We were unaware that such a shared standard does not exist for Grey Literature.

A by-product of having a shared XML standard for ingesting content is that there also exists a shared framework of understanding for online presentation of that content type. We failed to understand that the lack of a common DTD for Grey Literature would require us to not just develop Extensible Stylesheet Language (XSLT) to ingest the XML, but also we would have to define how that content would be presented to the end-user on the platform, requiring extensive User Experience (UX) and User Interface (UI) considerations and development.

#### *What have we learned?*

The biggest takeaway from the content preparation process is that each content type is unique and that uniqueness brings challenges and further questions. Although at a basic level all content types are some combination of words and pictures, they also come with a set of end-user expectations as to how that content should be consumed. Presenting a map, for example, in the same manner as we present a scholarly article could confuse the end user and ultimately lead the user to give up on trying to use that content type in our platform for their research.

Despite this reality, we also learned that the majority of content types share many metadata elements. Whether a piece of content is a conference proceeding, internal report, or continuing education article, that content has an author, a publication date, a title, and other common metadata elements. This realization meant that it would be possible to reuse large amounts of code for DTDs and XSLTs that would be created to support the content.

#### *What questions remain?*

Concerning content preparation there is one standout question that remains unanswered, is there a need for a shared DTD for Grey Literature?

The absence of a common, standardized DTD in the vein of JATS or BITS within the Grey Literature space makes content acquisition and ingestion into publishing platforms more complex and time-consuming than might be necessary. As part of this project, GeoScienceWorld spoke with various other organizations engaged in similar Grey Literature presentations to research communities and the absence of a common DTD forces them to do custom development projects that are costly and reflect the priorities of the organizations themselves rather than the user communities this content is intended to serve.

A secondary question to the above is does GeoScienceWorld take the lead in trying to define, create, and socialize such a DTD, and if so, with whom do we work to bring this to fruition?

## Implications for Search

### *What went well?*

When it comes to how end users search for content in a research platform, GeoScienceWorld understands clearly that it is the user that defines how they search. It is not for publishers or platform providers to force users into particular search experiences which limit the content on behalf of the user.

GeoScienceWorld understands the importance of the web best practice of allowing users to direct their own search experience. In having this clear understanding, GeoScienceWorld realized the need for visual cues as part of the search experience to help end users understand the results they were seeing in the browser.

The current GeoScienceWorld search experience includes such visual clues when it comes to the authorization state of a piece of content, showing the user if they have the authorization to view with a green check mark or a shopping cart icon if not. The search experience also uses visual cues to identify Open Access content through the use of an unlocked orange padlock. Knowing that such visual aids are useful to the user, GeoScienceWorld knows that a similar set of cues will be required to identify the peer-review status of the content. We are aware of sets of icons already available for this kind of cue and will implement them in our search results page when we have Grey Literature as part of the platform.

### *What could be better?*

While we have a very clear understanding that it is important to let the user direct their own search experience, we need to better understand the kinds of filters and facets that users find useful when refining their search.

At present the search results page has twelve facets, allowing the user to refine their search by content format, article type, publisher, publication date, and so on. GeoScienceWorld does have a reliable sense of which facets are regularly used, or even if the information and order of the facets are beneficial to the user. To better understand user behavior in general, GeoScienceWorld has implemented a tool called HotJar to create heat maps showing us where on the page users scroll and click. Our goal is that this information will allow us to present only the facet that researchers find useful.

### *What have we learned?*

A major learning for GeoScienceWorld as a result of this project has been to make real our understanding that the user must be free to direct their own search experience. While we were clear that this best practice was key to this particular project, we had allowed ourselves to fall into the trap of directing a user's search experience without reference to their goals in another area of the platform.

The feature in question was that when a user entered a search term in the quick search that was also a keyword in our thesaurus, the site redirected the user to the thesaurus term page for that keyword. This had the unintended consequence of users searching on common geoscience terms and being forced into a very limited set of search results.

GeoScienceWorld has subsequently removed this functionality from the platform and has received positive feedback on the scope of results being returned in the search.

As GeoScienceWorld has spent a lot of time as a result of the Grey Literature project investigating how search functions and is presented on our platform, we have realized that search results contain overwhelming amounts of metadata. Each search result presents the user with at least eight metadata items, as well as a snippet of the content, where the search term is found. The search result also includes options to view the abstract, download the PDF, purchase the content, or add to the Citation Manager tool. Based on feedback from users, we have realized that half the metadata fields displayed

are not relevant to the search itself, and instead serve to clutter the results and overwhelm the user.

#### *What questions remain?*

As we further consider the implications for search that this Grey Literature project has uncovered there are two questions that we are still seeking to answer.

Firstly, and touching on the amount of metadata displayed to the user within the search result, what metadata elements are of the most importance and value to the researcher to encourage them to investigate this piece of content further? It seems obvious that the content title is essential, but beyond that is there anything that is "sacred"? We are also investigating whether the true value of metadata is behind the scenes in the technology running the search functionality as opposed to being presented in the interface.

The second question that remains is when there exists a visual cue within the search result does the system need a corresponding facet to allow the user to filter results based on a visual cue? In our investigations we see very little evidence of user behavior to filter results where a visual cue exists, rather they recognize the cue and use that as a guide to investigating further, for example opening the abstract flyout from the search result.

### **New Business Models**

#### *What went well?*

As part of GeoScienceWorld's early investigations into this content set, we very quickly identified that the Grey Literature elements would be best suited to our industrial customer base, including consultants and other non-academic organizations in the geoscience space.

Having identified the primary customer group for this content we had to find business models that would suit organizations where the concept of an ongoing subscription to a body of content is not relevant. As such, we decided that implementing a tokenized purchasing model would be the most suitable approach. This product offering allows an organization to buy a bundle of tokens that can be redeemed against content as needed. The greater the number of tokens purchased, the more cost-effective the product is on a per-article basis.

GeoScienceWorld developed and launched this token approach specifically for the corporate market early in 2022. Although the research platform currently only contains traditional journal article and ebook content, we built tokens in such a way that it is scalable to new content types with very little further development.

#### *What could be better?*

Although we readily had a sense of the primary market for the Grey Literature subset of content in the project, we had a slightly blinkered view of secondary markets, being limited in our thinking of available markets as being either academic or corporate. At a recent conference, I attended a session where a publisher discovered a wide range of organizations that neither fell into the academic nor corporate camps by virtue of their IP ranges being held in the IP Registry, a global database of IP addresses.

The publisher in question wanted to understand what organizations were using their Open Access content, therefore showing a general interest in the publisher's content. They discovered over 20 identifiable organizations, mainly NGOs and governmental organizations that were regularly using their content and thus presented additional sales opportunities. GeoScienceWorld is also looking at working with the IP Registry to try and identify similar organizations who show an interest in the content we host, and could thus open new markets to us.

Despite identifying our primary market, and having a method for potentially uncovering unexpected secondary markets, GeoScienceWorld is keen to find a unique selling point

for our academic customers, who form the majority of our customer base. We are convinced that this Grey Literature set has value to the researchers working in the academic institutions that we serve.

#### *What have we learned?*

It was in the realm of new business models that we had our most surprising learnings. In conversations with governmental leaders within the geoscience space, specifically several state geologists, we learned that organizations are willing to pay for access to content that is nominally free. Such organizations find value in the content being collated into a single online repository that they pay for access to have benefited from the convenience of this kind of content being readily discoverable in a centralized location.

This particular revelation was startling to GeoScienceWorld as the tools surrounding the content on our research platform had always been a secondary consideration to the content itself. Discovering that purchasers find monetary value in the features and functionality and of course the convenience, of having paywalled and free content in a single platform.

We also learned that our end users, the research community as opposed to the librarian community, are less concerned about the peer-reviewed status of the content than they are about the value and reproducibility of the science behind the article itself. We suspect that this may be a behavior specific to researchers within the geosciences, and are seeking to validate this assumption.

#### *What questions remain?*

While having identified a tokenized purchasing model as key to reaching our primary, corporate, market, we are left asking the question, what business models would suit potential secondary markets uncovered by researching organizations in the IP Registry using GeoScienceWorld's Open Access and free content regularly? Are there opportunities to increase our subscription sales within these secondary market organizations? Would a "freemium" model, where a user has restricted access to the content and tools on the platform be worth investigating further? A third approach we are looking into, especially with public access content is to have users pay a small subscription fee to access advanced search and content manipulation tools on the platform, whilst keeping the content itself free to view.

GeoScienceWorld is also investigating whether the data supporting the content, whether it be spreadsheets, code, or other non-textual assets is in some ways more valuable to the research community than the article or ebook chapter?

### **Conclusions**

In my original paper, written in support of the presentation I gave in Prague, I wrote that as

"GeoScienceWorld embarks on the next phase of this migration project, actually building out the features needed to support grey literature, we expect most of our assumptions to be challenged, the requirements to need changing many times, and to have a strong partnership with our platform provider to meet the architectural problems that will likely pop up as we try to make grey literature work in a framework specifically designed for scholarly content".

It is fair to say that the assumptions we made in the initial phase of this migration project have been challenged, possibly more so than we ever imagined. However, as an organization committed to implementing an Agile mindset, we embrace that challenge, especially as revealed through this retrospective process.

We are convinced that Grey Literature has very real, tangible, value to the research community that uses the traditional peer-reviewed literature that forms the basis of our

platform. That value makes Grey Literature repositories an attractive proposition in our content acquisition efforts.

Although the project has effectively been in stasis since the beginning of the pandemic, GeoScienceWorld is confident that we are ready, able, and willing to pick up this project and see it through to completion. With the many things we have learned and developed in response to this migration project, we feel that Grey Literature is an important part of our future.

Much of that confidence comes from GeoScienceWorld's efforts to implement an Agile mindset within the organization, and adopting key Agile ceremonies such as retrospectives allows us to have a firm basis for moving forward. Agile also allows us to learn and improve from projects that might otherwise be regarded as a "failure". We have been able to implement our learnings from this process in other aspects of our organization and the work we do.

The Agile mindset that we are seeking to make the bedrock of our organizational culture allows us to view all project outcomes as opportunities for continuous improvement. Our aim as an organization is not to accept the status quo, but to always be improving our workflows and processes to add value to our various customer groups, whether that be librarians, corporate geoscientists, or the academic research community.

Perhaps our biggest conclusion as a result of this project so far, and the retrospective we undertook, is that there is a distinct need for leadership within the Grey Literature community when it comes to metadata preparation for content loading and presentation standards. As more and more traditionally scholarly publishers bring non-journal and book chapter content into the platforms that we already use, there is an opportunity for the Grey Literature community to work together with the platforms to define shared standards to make Grey Literature a more accepted part of the research databases that science relies upon.

## References

The Challenges of Incorporating Grey Literature into a Scholarly Platform [online]: 2019, National Repository of Grey Literature, National Library of Technology, Prague, Czechia [Accessed on 18th November 2022]. Available from: <https://invenio.nusl.cz/record/407833>

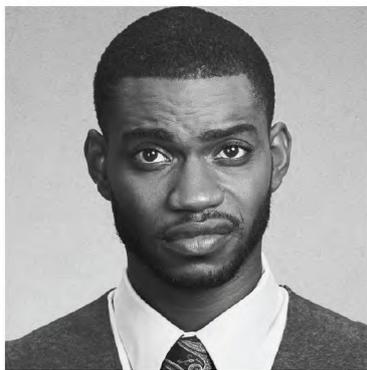
Agile Manifesto [online]: 2001, The Agile Alliance [Accessed on 18th November 2022]. Available from: <https://agilemanifesto.org/>

The IP Registry Global IP Address Database [online]: 2022, PSI Ltd, Oxford, UK [Accessed on 18th November 2022]. Available from <https://theipregistry.org/>



## 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**

## Information Fatigue Syndrome and Digital Burnout

Dobrica Savić,

IAEA Consultant in Nuclear Knowledge Management

Video Presentation

<https://av.tib.eu/media/59863>

### Abstract

*Around 2000 years ago, the Roman philosopher Seneca the Younger complained that his peers were wasting their time and money accumulating too many books, admonishing that "the abundance of books is a distraction." Instead, Seneca recommended focusing on fewer but better-quality books and reading them thoroughly and repeatedly. After the invention of the printing press in the 15<sup>th</sup> century, the plethora of information the machine produced was thought by some to be distracting.*

*Today, information overload and digital over-stimulation lead to digital burnout — a situation where physical and mental exhaustion is caused by spending too much time in front of screens. Symptoms of this condition are apathy, indifference, or mental exhaustion arising from exposure to too much information. Stress induced by attempts to assimilate excessive amounts of information from the media, particularly social media, the internet, or work makes us ill by interfering with our sleep, sabotaging our concentration and undermining our immune system. David Lewis, a British psychologist, calls this ailment Information Fatigue Syndrome (IFS).*

*This paper looks at Information Fatigue Syndrome (IFS), and particularly at the contribution of grey literature (GL) to IFS. To better understand the amount of grey literature around us, it is enough just to look at the GreyNet website which lists over 150 grey literature types. They include articles, blogs, images, videos, emails, web pages, press releases, lectures, manuals, academic theses, and many others. Today, grey literature occupies a considerable role and is one of the main contributors to our digital burnout. The diversity of GL types, volume generated and shared, frequency of creation and change, currency, veracity, and value will be reviewed.*

*The purpose of the paper is to generate suggestions on how to successfully deal with Information Fatigue Syndrome. Additionally, it will offer suggestions on how to continue using valuable grey literature efficiently, and how to do all of this without causing unnecessary stress or wasting time.*

**Keywords:** *digital burnout, grey literature, Information Fatigue Syndrome*

### Introduction

Is there anyone among us today that has not experienced the nagging feeling of having too much information and too little time to deal with it? Do you sometimes feel mental exhaustion from being exposed to too much information? You overly multitask but your concentration and memory fade, while your irritability grows. Your feeling of helplessness grows, together with relationship problems with your colleagues and with your loved ones at home.

Welcome to digital burnout and to Information Fatigue Syndrome where the overwhelming amount of grey literature plays a significant role.

We believe that information fatigue is a new phenomenon, something *sui generis* to the 21<sup>st</sup> century. However, around 2000 years ago, the Roman philosopher Seneca the Younger complained that his peers were wasting their time and money accumulating too many books, admonishing that "the abundance of books is a distraction" (AZ Quotes).

Seneca recommended focusing on fewer but better quality books and reading them thoroughly and repeatedly.

After Johannes Gutenberg invented the printing press in the 15th century, the plethora of information produced, shared and disseminated widely was thought to be very distracting and counterproductive.



Similarly, the introduction of personal computers and especially the creation of the World Wide Web has dramatically increased the amount of information easily available to all of us.

### Information Fatigue Syndrome (IFS)

It is well-established that information overload and digital over-stimulation cause digital burnout — a situation where physical and mental exhaustion is caused by spending too much time in front of screens.

Stress induced by attempts to assimilate excessive amounts of information from the media, particularly social media, the internet, or work, makes us ill by interfering with our sleep, sabotaging our concentration and memory, and undermining our immune system and overall well-being.

Information Fatigue Syndrome is defined as a weariness or overwhelming feeling of being faced with an indigestible or incomprehensible amount of information.

If we look at the term information fatigue syndrome, also known as information overload or information intoxication, we come across David Lewis (Wikipedia), a British psychologist, who lived from 1941 to 2001, and who is credited with coining the first use of the term.

He said that "Having too much information can be as dangerous as having too little. Among other problems, it can lead to a paralysis of analysis, making it far harder to find the right solutions or make the best decisions."



In his report, diabolically entitled *Dying for Information?* (Waddington, 1998) Lewis said in 1996 that an excess of information is strangling many businesses and causing mental anguish and even physical illness in managers at all levels. Lewis speculated that the problem would only worsen, and it seems that his prediction was correct.

His conclusions came from a Reuters survey of 1,300 business people in Britain, the US, Singapore, Hong Kong and Australia, which included junior, middle and senior managers in a variety of industry sectors. Two-thirds of those interviewed indicated that stress, attributed to dealing with too much information, had damaged their personal relationships, increased tension with colleagues at work, and contributed to a decline in job satisfaction.

More than 40% felt that important decisions were delayed and the ability to make choices was hampered by excess information. The cost of collecting the surplus data exceeded its value. One-third said they suffered from health problems as a direct consequence of stress related to information overload.

### Grey Literature Fatigue Syndrome (GLFS)

Let's look at grey literature and the role it plays as a part of information fatigue syndrome. It is believed that grey literature<sup>1</sup> overload is a major part of information overload and a cause of information fatigue syndrome.

<sup>1</sup> Grey literature represents any recorded, referable and sustainable data or information resource of current or future value, made publicly available without a traditional peer-review process (Savic 2017).

There are two major factors that make grey literature the main contributor to our information fatigue syndrome. They are an extensive variety of grey literature document types and extremely high amounts, the quantity of grey literature output.

Also, grey literature is highly contextual and often software dependent, so it is hard to collect and process, and even harder to make sense of and preserve for future use.

Examples are social media, news items, emails, reports, and data.

### **Grey Literature Types**

To illustrate the huge number of grey literature types, we can consult the GreyNet website (GreyNet, 2022). It lists over 150 document types including databases, data sets, data sheets, data papers, satellite data, census data, and product data, just to mention some of the many data types. See Appendix 1.

### **Information Fatigue Symptom**

So, what are the major symptoms, the visible and observable behavioural and other changes characteristic of someone who suffers from information, including grey literature fatigue syndrome?

They are:

- Apathy
- Indifference
- Mental exhaustion arising from exposure to too much information
- Poor concentration
- Short-term memory failure
- Overly multitasking, resulting in incomplete tasks
- Over-stimulation causing headaches and nausea
- Tension
- Relationship problems at home
- Occasional irritability
- Frequent feelings of helplessness
- Compulsive need to be connected to the internet

### **Main Causes of Information Fatigue Syndrome**

As the great novelist and poet Gertrude Stein put it - *Everybody gets so much information all day long that they lose their common sense.*

Having established the starting ground for information and grey literature fatigue syndrome, let's look now at its main causes.

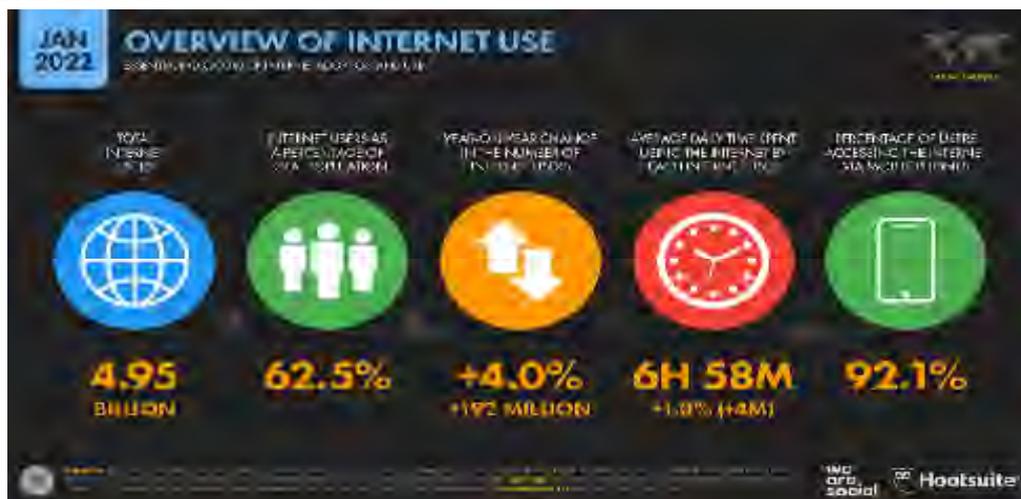
#### **A. Digital world**

Since we live in a different world today, the digital world, its main characteristics and attributes make up the basis for digital burnout.

As of January 2022, the total population of the world was 7.91 billion. 57% of those live in urban areas, and over 67% of the total population or 5.31 billion use a mobile phone.

4.95 billion people, which is well over 60% of the population, are internet users, and almost all of the people connected to the internet are also active social media users.

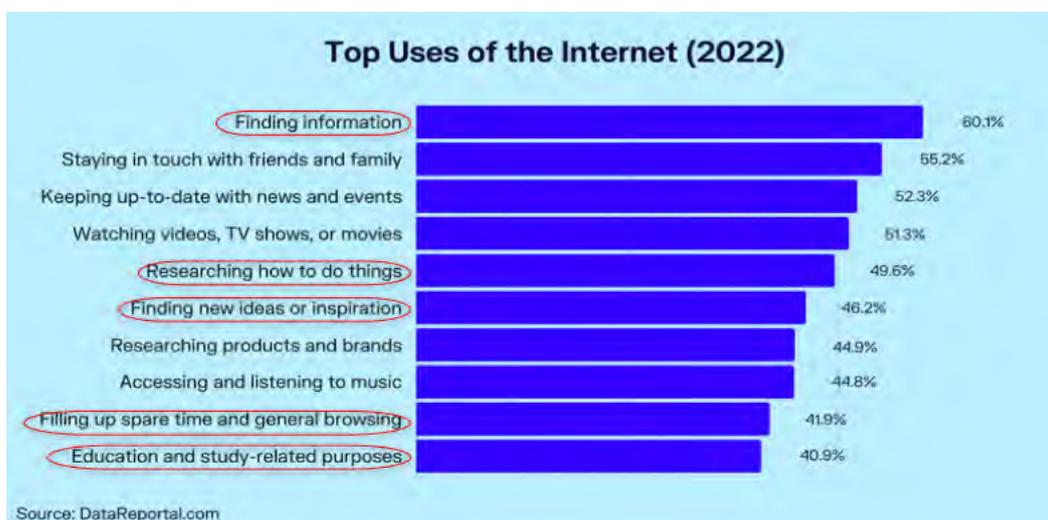
## B. Internet use



Consulting the illustration above (Datareportal, 2022) regarding the overview of internet use, let's concentrate on just the two last figures.

The average DAILY time spent using the internet by each internet user is almost 7 hours. This is an astonishingly high number of hours spent, which should make us all think hard about its usefulness and our reasons for spending so much time using it.

For all of us who are creating, providing, and organizing information and grey literature on the internet, the statistic that 92.1% of users access the internet through their mobile phones should be of the utmost importance. The previous reality of personal computers (PCs) occupying this major role is not valid anymore. It is clear that focus has dramatically shifted towards the use of mobile devices. To that end, we need to direct our efforts towards this relatively new tool and adjust our internet and web presence consequently.



The above graph (Oberlo 2022) of top internet uses shows that the majority of people are using it to search for information, communication, and entertainment. It is encouraging that there is a high percentage of use geared towards education and study, which should be encouraged, especially for school-age internet users.

## C. Social media

How much time do we spend on social media apps? Statistics show that monthly, on average, we spend almost 24 hours on YouTube, 20 hours on TikTok, and the same on Facebook. Another 19 hours are spent monthly on WhatsApp, the most popular communication app.

Generally speaking, an average social media user spends between 80-90 hours monthly on various social media apps. Translated into weekly work hours, we spend over two work weeks browsing various social media or reading messages we've received.

So let's review this, on average, every user spends almost half of his or her productive work hours on social media blindly devouring content that is often questionable.

Let's have a quick look at some YouTube statistics in 2022 (DemandSage 2022):

- There are 2.6 billion YouTube users
- Viewers watch over 1 billion hours of video every day
- Localized in over 100 countries and 80 languages
- 63% of watch time derives from mobiles
- 400 hours of video uploaded every minute
- The most popular video platform



#### **D. Data**

The amount of data available around the world in 2020 was estimated at 59 zetabytes (ZB). While it is predicted that this will reach a mind-boggling 175 ZB by 2025.

One ZB (Intellobics 2011) is equivalent to one trillion gigabytes. If each bit is a coin around 3mm thick, one ZB made up of a stack of coins would be 2,550 lightyears. This would get you to our nearest star system, Alpha Centauri, 600 times (TechCentral 2021). To put it into a closer perspective, one ZB is equivalent to 36,000 years' worth of HD-quality video.

#### **E. Email**

In a work environment the quantity of emails sent and received represents probably the main contributor and cause of information fatigue syndrome. It is estimated that there are 4.3 billion email users around the world (Oberlo 2022), and there is also an evident trend that the number of email users is growing.

People around the world send over 333 billion emails daily (Statista 2022).

#### **F. Journals**

Regarding journals, in 2009 we passed the 50 million mark of the total number of scientific papers published in various journals since 1665 (Jinha 2010). Approximately 2.5 million new scientific papers are published each year.

As of 2014, there were approximately 28,100 active scholarly peer-reviewed journals (Ware 2015). This excludes the increasing number of predatory, fake scientific journals, which produce high volumes of poor-quality research.

#### **G. Books**

It is estimated that there are between 500,000 and one million books published annually (Bobby 2022).

With self-published authors, there are close to 4 million new book titles published each year. Unfortunately, the typical self-published author sells only about five copies of his/her book.

The average US book now sells fewer than 200 copies per year and fewer than 1000 copies over its lifetime.

According to Google, there have been 130 million books published since the invention of Gutenberg's printing press in 1440. However, this doesn't factor in books published after 2010, nor does it include self-published book titles.

In 2021 a total of 826 million books were sold in the US. An interesting fact is that 75% of people surveyed in the US prefer print to e-books or audiobooks.

Let's remind ourselves of Seneca's message, that it does not matter how many books we have, but rather how good these books are.

## Information Fatigue Syndrome Solution

Is there a solution to Information Fatigue Syndrome?

The solution to Information Fatigue Syndrome is not simple or straightforward.

Some main actions to deal with it are the following:

- Filter all the information that comes your way and make sure that you don't contribute to the digital burnout of others by spreading and sharing unnecessary information
- Sharpen your focus when looking for and using grey literature and any other information
- Focus on essential, not on interesting
- Prioritize
- Pick reliable and trustworthy sources of grey literature
- Delegate
- Ask for help
- Learn to say NO
- Shut down disrupting devices
- Separate business from private time
- Relax, go for walks, meditate

There is also a role that information specialists could play. They can:

- Help with filtering information
- Maintain lists of high-impact resources
- Prioritize readings and research materials
- Do preliminary search
- Determine reliable and trustworthy sources
- Offer learning and research hubs
- Provide opportunities to experience modern IT environments
- Offer information and KM training
- Encourage exploration, creation, and collaboration
- Provide no-stress and quiet spaces
- Become pillars of open-access and open science

And finally, let's consider a huge wastepaper basket to be a possible solution for information fatigue syndrome, since as Albert Einstein said, '**information is not knowledge!**'

## References

AZ Quotes. <https://www.azquotes.com/quote/1261725>

Bobby. 2022. The number of books published each year. <https://www.tagari.com/the-number-of-books-published-each-year/>

Datareportal. 2022. <https://datareportal.com/reports/digital-2022-global-overview-report>

DemandSage. 2022. YouTube Statistics (2023) — Trending Facts & Figures Shared! <https://www.demandsage.com/youtube-stats/>

GreyNet. 2022. Document Types in Grey Literature. <http://greynet.org/greysourceindex/documenttypes.html>

Intellobics. 2011. Understanding information measures. <https://intellobics.com/2011/02/14/understanding-information-measures/>

Jinha, Arif E. 2010. Article 50 million: An estimate of the number of scholarly articles in existence. July 2010 *Learned Publishing* 23(3):258-263. <https://onlinelibrary.wiley.com/doi/abs/10.1087/20100308>

Oberlo. 2022. Top Uses of the Internet in 2022. <https://www.oberlo.ca/statistics/what-is-the-internet-used-for>

Savić, D., 2017. Rethinking the Role of GL in the Fourth Industrial Revolution. 10<sup>th</sup> Conference on GL and Repositories: proceedings [online]. Prague: National Library of Technology. <http://nrgl.techlib.cz/index.php/Proceedings>. ISSN 2336-5021. Also published by TGJ (The Grey Journal) Special Winter Issue, Volume 14, 2018.

Statista. 2022. Number of sent and received e-mails per day worldwide from 2017 to 2025. <https://www.statista.com/statistics/456500/daily-number-of-e-mails-worldwide/>

TechCentral. 2021. The world's data: The world's data: How much we're producing and where it's all stored. <https://bit.ly/3WX41q6>

Waddington, Paul. 1998. Dying for Information? A Report on the Effects of Information Overload in the UK and Worldwide. <http://www.ukoln.ac.uk/services/papers/bl/blri078/content/repor~13.htm>

Ware M, Mabe M. 2015. The STM report. The STM Report: An overview of scientific and scholarly journal publishing Celebrating the 350th anniversary of journal publishing. <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1008&context=scholcom>.

Wikipedia. [https://en.wikipedia.org/wiki/David\\_Lewis\\_\(philosopher\)](https://en.wikipedia.org/wiki/David_Lewis_(philosopher))

## Appendix 1: Grey Literature Types

<p><b>A</b> Abstracts Advertorials Announcements Annuals Article</p> <p><b>B</b> Bibliographies Blogs Booklets Brochures Bulletin Boards Bulletins</p> <p><b>C</b> Call for Papers Case Studies Catalogues Chronicles Clinical Trial: - Source Document Codebooks Conference Papers Conference Posters Conference Proceedings Country Profiles Course Materials</p> <p><b>D</b> Databases Data Papers Datasets Datashets Deposited Papers Directories Discussion Papers Dissertations Doctoral Theses</p> <p><b>E</b> E-Prints E-texts Enhanced Publications Essays ETD (Electronic Theses and Dissertations) Exchange Agreements</p> <p><b>F</b> Fact Sheets Feasibility Studies Flyers Folders Forum: - Internet</p> <p><b>G</b> Glossaries Government Documents Green Papers Guidebooks</p> <p><b>H</b> Handbooks House Journals</p>	<p><b>I</b> Image Directories Inaugural Lectures Indexes Interactive Posters Internet Reviews Interviews</p> <p><b>J</b> Journals: - Articles - Grey Journals - In-house Journals - Non-commercial Journals - Synopsis Journals</p> <p><b>K</b> K-blogs</p> <p><b>L</b> Leaflets Lectures Legal documents Legislation LibGuides</p> <p><b>M</b> Manuals Memoranda</p> <p><b>N</b> Newsgroups Newsletters Notebooks</p> <p><b>O</b> Off-prints Orations</p> <p><b>P</b> Pamphlets Papers: - Call for Papers - Conference Papers - Deposited Papers - Discussion Papers - Green Papers - White Papers - Working Papers Patents Policy Documents Policy Statements Posters Précis Articles Preprints Press Releases Proceedings Product Data Programs Project: - Deliverables - Information Document (PID) - Proposals - Work Packages - Work Programmes</p> <p><b>Q</b> Questionnaires</p>	<p><b>R</b> Readers Registers Reports: - Activity Reports - Annual Reports - Bank Reports - Business Reports - Committee Reports - Compliance Reports - Country Reports - Draft Reports - Feasibility Reports - Government Reports - Intelligence Reports - Internal Reports - Official Reports - Policy Reports - Progress Reports - Regulatory Reports - Site Reports - Stockbroker Reports - Technical Reports Reprints Research Memoranda Research Notes Research Proposals Research Registers Research Reports Reviews Risk Analyses</p> <p><b>S</b> Satellite Data Scientific Protocols Scientific Visualizations Show cards Software Specifications Speeches Standards State of the Art Statistical Surveys Statistics Supplements Survey Results Syllabus</p> <p><b>T</b> Technical Documentation Technical Notes Tenders Theses Timelines Trade Directories Translations Treatises Tutorials</p> <p><b>W</b> Website Reviews WebPages Websites White Books White Papers Working Documents Working Papers</p> <p><b>Y</b> Yearbooks</p>
---	--	---

Source: <https://greynet.org/greysourceindex/documenttypes.html>

## Open Science as an Opportunity for Academic Grey Literature – A Systematic Review

Joachim Schöpfel, University of Lille, ULR 4073 - GERiiCO - Groupe d'Études et de  
Recherche Interdisciplinaire en Information et Communication

Hélène Prost, CNRS, ULR 4073 - GERiiCO - Groupe d'Études et de Recherche Interdisciplinaire en  
Information et Communication, Lille, France

### Video Presentation

<https://av.tib.eu/media/58582>

### Abstract

*What is the future of grey literature? Is open science an opportunity for grey literature (more grey literature, more visibility and impact of grey literature), just another challenge (issues that need awareness and further action) or even a threat (less grey documents, less impact)? The following paper presents findings from a systematic review of recent studies on grey literature and open science; it includes recent reports and initiatives, and builds on our own former empirical research work.*

*The findings of the review are presented in four sections: the concept of bibliodiversity (production); the development of open repositories (dissemination); the transformation of research assessment (evaluation); the application of FAIR principles (processing). Some leading questions: Which is the common part of the concepts of bibliodiversity (Jussieu Call) and grey literature, and what does this mean for the future of grey literature? Which are the issues of grey literature in open repositories, and how do repositories impact the dissemination of grey literature? How do the recent initiatives for a new system of research assessment affect the grey literature (San Francisco Declaration DORA, European Commission Scoping report, OSEC Call of Paris)? Are the FAIR principles relevant for the processing of grey literature, and if so, which ones and in which way?*

*The paper is a scientific contribution to the analysis of the development of grey literature in the academic research environment of the 21st century. It is a rejoinder of the 2021 panel discussion on the "Next Generation Grey"; it tries to sum up what is known about grey literature and open science, and it makes some recommendations for grey literature producers, service providers and repository managers.*

*Keywords: Open science, open access, open repositories, grey literature*

### Introduction

What is the future of academic grey literature? A couple of papers have made assumptions in the past about the perspectives of grey literature, for instance, regarding the impact of new technologies or of research policies. Some predicted an unprecedented increase of the volume of grey literature, with global and unrestricted accessibility (Mackenzie Owen, 1997), while others did not observe an unusual growth of grey documents (Artus, 2005) or projected a greater diversity of grey literature and convergence between grey and white (Schöpfel, 2006). In GreyNet's 2021 online survey on the future of grey literature, two-third of the respondents among GreyNet's own community of practice agreed that "*in academic institutions, the affirmation of open science and open access principles significantly favor the production, publication, and retrieval of grey literature*" (Schöpfel et al., 2022). It remains uncertain, however, what this exactly means, and if open science is an opportunity or a threat for grey literature, or just another challenge that needs awareness and further action, or all together (table 1).

Scenario	Examples
Opportunity	More grey documents More diversity Better visibility More impact
Challenge	Curation of metadata Persistent identifiers Openness
Threat	Less grey documents Less impact

Table 1. Scenarios for grey literature

As for grey literature in repositories, one respondent wondered “(...) if it is still grey literature (because) published grey literature is no more grey”. The boundary between white and grey is still a matter of discussion, such as the distinction between data and documents. Probably it is the very concept of grey literature that is at stake, being often, especially in systematic reviews, described as “unpublished” and “non-reviewed research” (Schöpfel & Prost, 2021) which is not exactly its meaning in library and information sciences, where the focus is on dissemination and acquisition (Schöpfel & Farace, 2018).

Forecasting the future of grey literature may appear hazardous in an environment of constant and rapid change. Our paper takes a somewhat different approach, insofar we try to find out how (and if) some essential topics of the actual debate on open science have been linked to the field of grey literature. Based on a systematic review, we assessed the topics of bibliodiversity (production), of open repositories (dissemination), of research assessment (evaluation) and of FAIR principles (processing). Some results confirm what we already know; others are more surprising.

The paper is a contribution to the analysis of the development of grey literature in the academic research environment of the 21<sup>st</sup> century. It is also a rejoinder to GreyNet’s panel discussion on the “Next Generation Grey” in 2021. We try to sum up what is known about grey literature and open science and to make clear what is questionable and what needs more insight.

### Methodology

We conducted a systematic review on relevant studies published between 2018 and 2022 on grey literature and open science in two bibliographic databases (Web of Science, Scopus), in the Grey Guide repository and with two academic discovery tools (Dimensions, Google Scholar). The review was conducted on August 2, 2022. The retrieval strategies and filters are described in the appendix. Diagram 1 shows the results of the systematic review.

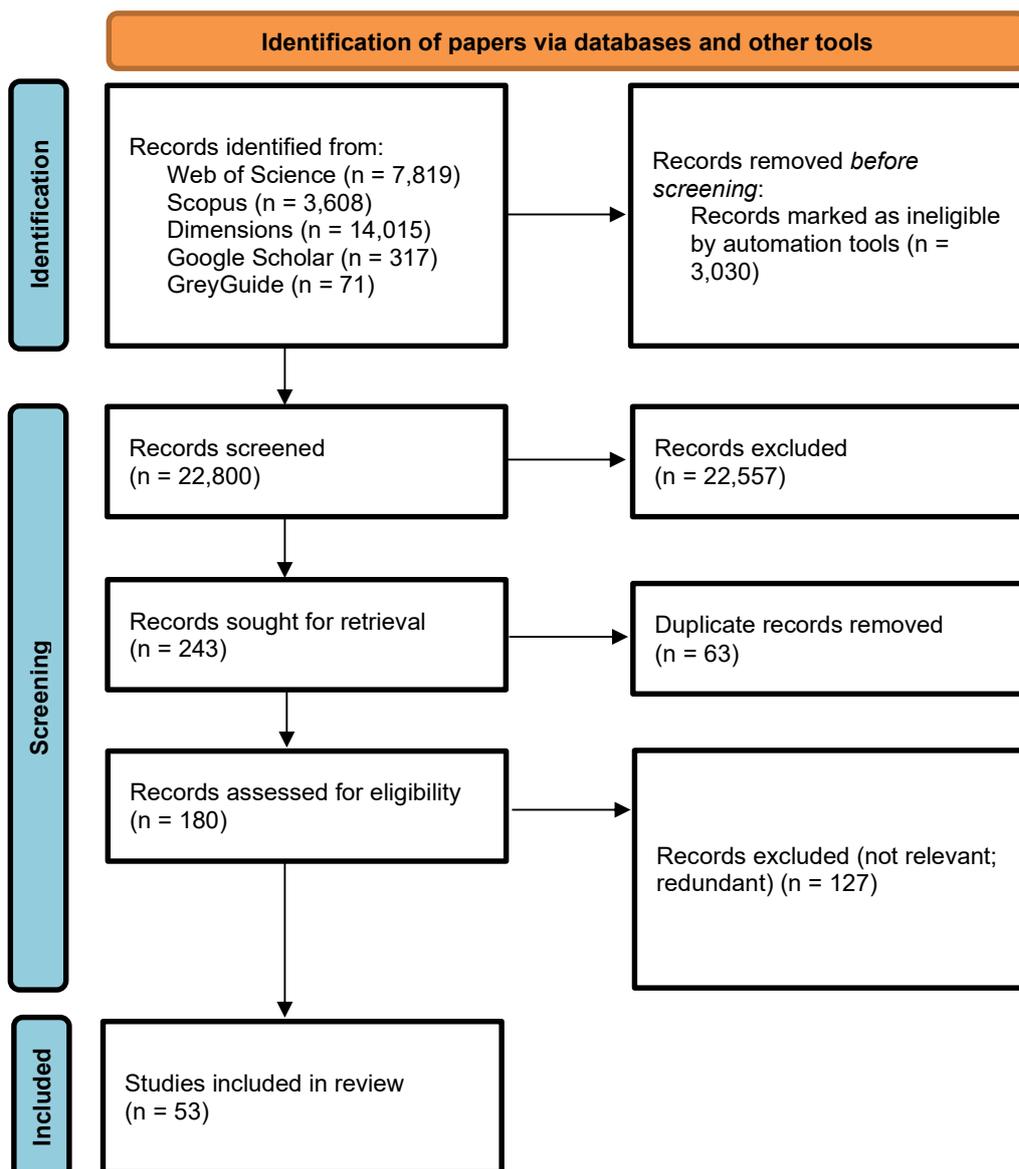


Diagram 1. PRISMA 2020 flow diagram of the systematic review on grey literature

We were looking on recent, original research work on the impact of open science on grey literature, with relevant insights on the production and dissemination of grey literature, on its role for research assessment and on the application of FAIR principles. Initially, we identified 25,580 records. Most records have been discarded because they mentioned grey literature only as part of systematic reviews; others have been excluded because they were not original research or not relevant for open science. Half of the 53 included studies are papers presented at the *International Conferences on Grey Literature* and/or published by GreyNet's *The Grey Journal*, i.e., products of the GreyNet's community of practice.

## Results

### The call for bibliodiversity (production)

Open access to scientific information is a main pillar of open science. There is no doubt that grey literature plays a significant role in the development of open access (OA), as many repositories contain grey literature (figure 1) where it represents up to 30% of the content (Schöpfel et al., 2020).

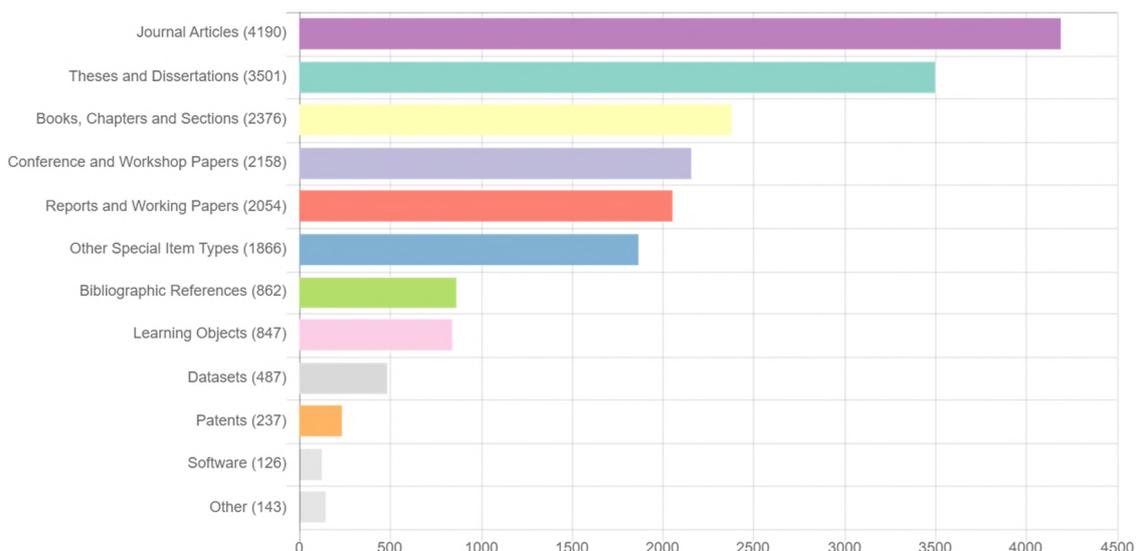


Figure 1. Content types overview in open repositories (N=5,971)  
(source: OpenDOAR, October 23, 2022)

However, the academic and political debate on open access mainly concerns journals and seems to ignore grey literature, as the main challenge is the business model of OA journal publishing and its economic and political impact. In the usual conceptualization of OA publishing, grey literature is more or less invisible, somewhere in the small field of non-peer reviewed items, free for readers and for authors, protected by copyright, along with preprints (figure 2).

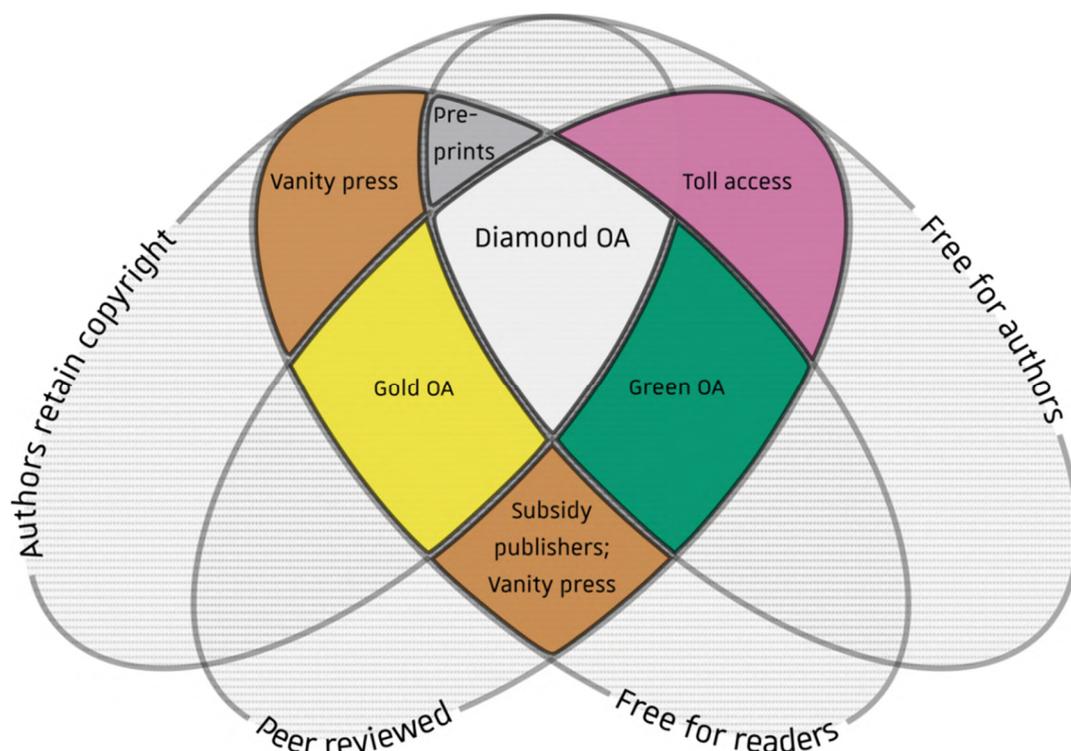


Figure 2. Different types of open access in scholarly publishing  
(source: Wikipedia Open Access<sup>1</sup>)

<sup>1</sup> By Jamie Farquharson - <https://doi.org/10.6084/m9.figshare.6900566.v1>, CC BY 4.0, <https://commons.wikimedia.org/w/index.php?curid=111532339> )

Recently, a couple of initiatives have been taken against the development of a dominant OA publishing model based on article processing charges (APC). While some focus on journals, such as the *Action Plan for Diamond Open Access*<sup>2</sup> promoted by Science Europe and the cOAlitionS, others are more open and inclusive, like the *Jussieu Call for Open Science and Bibliodiversity*<sup>3</sup> launched in 2017 by French scientists and professionals and promoting open access to scientific publications that encourages bibliodiversity and innovation and does not imply the exclusive transfer of subscriptions to APCs.

Even if the main concern of the *Jussieu Call* is journal and book publishing, its scope is larger as it explicitly invites for innovation and experimentation in the field of dissemination of research results and academic publishing. Our systematic review retrieved recent papers that highlight the richness and the diversity of unconventional academic literature, (partly) non reviewed and (partly) published outside of the usual commercial channels:

Preprints (non-reviewed, unpublished literature): Drawing on examples from the U.S. National Library of Medicine, Sheehan (2021) recommend the usage of elements of grey literature (preprints, study protocols, clinical trials...) to support open science as part of the response to the COVID-19 pandemic (figure 3). Even if some preprints will never become a published article, their dissemination via preprint servers and repositories contributes to the knowledge of the field, including deadlocks (Lombardi, 2021).

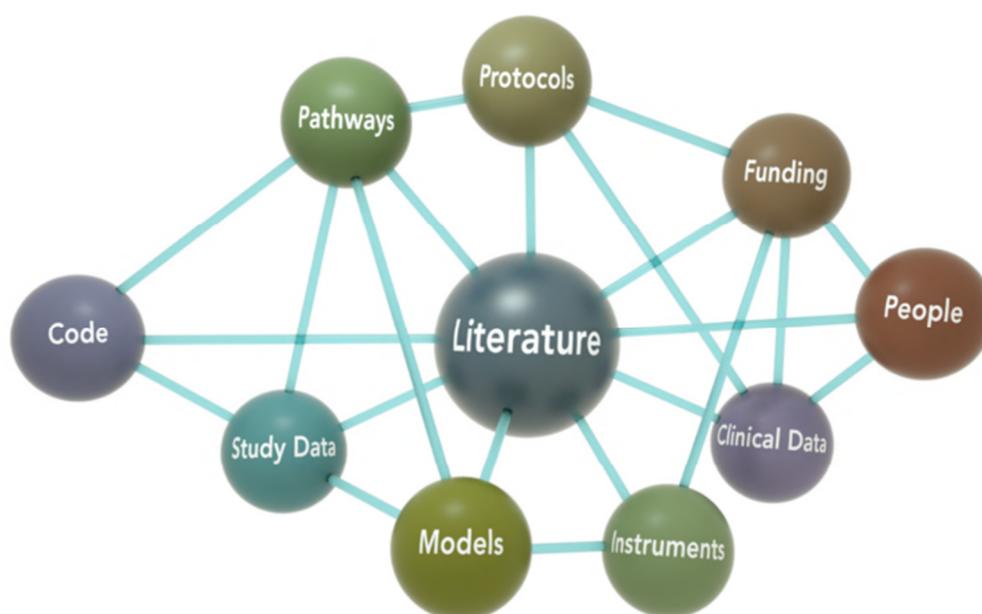


Figure 3. Network of interlinked research objects (source: Sheehan, 2021)

The Japan Atomic Energy Agency (JAEA) brings up the issues of long-term preservation and permanent access of conference proceedings and technical reports published in electronic media such as CD/DVD-ROM and flash memory (Kumazaki et al., 2019). Providing long-term preservation and improving discoverability of reports is also the purpose at Columbia University Libraries (Bielskas et al., 2022); in a similar approach, Cirkovic et al. (2019) argue for digitization, curation, preservation, and dissemination via open repositories of scientific heritage material, like older theses and dissertations.

<sup>2</sup> Action Plan for Diamond Open Access <https://www.scienceurope.org/our-resources/action-plan-for-diamond-open-access/>

<sup>3</sup> Jussieu call for Open Science and Bibliodiversity <https://jussieucall.org/>

Other categories and elements of grey literature have been put forward, in different domains:

- Law: Williams (2020) discusses the place of grey literature in the realm of legal materials (primary and secondary sources) and the challenges of discoverability of various types of grey legal materials.
- Linguistics: A study from the Italian CNR assesses the presence of a large variety of grey literature in the CLARIN Virtual Language Observatory, including translations, agreements, press releases, notebooks, and bulletins (Goggi et al., 2017)<sup>4</sup>.
- Open educational resources (OER): Puccinelli et al. (2020) provide a framework for the management of OER characterized by “*openness, flexibility, innovative approaches, digital dimension, liquidity and high granularity*” and insist on the development of “*validated and certified platforms, for providing quality levels in shared contexts, and persistent identification systems, for guaranteeing resources integrity (and) traceability*”.
- Another paper describes grey literature outside of academy, as extra-academic and/or professional documents, such as expert, CRM, R&D or corporate annual reports, market studies, executive directives, memoranda, white papers, and hearings, with examples in the fields of nuclear energy, aeronautics, archaeology and library and information sciences and four major challenges, i.e., monitoring, standardization, data, and preservation (Schöpfel, 2019).

Assuming that “*traditional publishing streams of books and journals capture just a fraction of the content (...) contextualized in open science*”, Gelfand & Lin (2020) provide a taxonomy of new forms of scholarship including theses, preprints, lab manuals, interviews, and other multimedia materials, and demonstrate how, in the context of open science, the research process contribute to “*outputs, innovation and new forms of grey literature*”. A similar, large range of resources has been identified by a scientometric study on systematic reviews and other articles making use of grey literature (Schöpfel & Prost, 2021).

Obviously, there is a common ground where both concepts, bibliodiversity and grey literature, overlap. As the *Jussieu Call* admits, the “*means to achieve the goal of Open Access are yet to be discussed*”. Our suggestion is that the signing institutions - research organizations and scientific communities, scientific and technical information professional associations and organizations, learned societies, scientific journals and editors – explicitly consider grey literature when they refocus “*the issue of business models (...) in the broader perspective of the editorial processes and methods upon which research and innovation will rely in the future*”, for the benefit of a “*very broad bibliodiversity*”.

The GreyNet community of practice, on the other side, should seize the opportunity of the Open Science initiatives for bibliodiversity, such as the *Jussieu Call*, for advocacy and promotion of the richness and diversity of grey literature which are in constant need for attention, curation, funding, and innovation, in order to increase its findability and accessibility.

### **The development of open repositories (dissemination)**

More than ten years ago, Daniela Luzi reviewed the development of institutional repositories and stated that “*grey literature is at home in open archives*” (Luzi, 2010). Which are the issues of grey literature in open repositories today, and how do repositories impact the dissemination of grey literature?

Like Daniela Luzi, most retrieved papers express the conviction, based on evidence, that repositories are good for grey literature and that grey literature, as an essential part of scientific output, should be part of repositories. “*Archiving in a repository is the way to let grey literature to become open*” (Lombardi, 2021). Based on a study on grey literature in

---

<sup>4</sup> This paper was retrieved because the communication was published in the conference proceedings in 2018.

repositories of Balkan countries, an invited paper of the 2018 UNESCO International Conference on Digital Presentation and Preservation of Cultural and Scientific Heritage recommends that universities and academic libraries “*should direct their efforts towards the creation of institutional repositories, allowing the preservation of and free access to the results of scientific studies performed at the institution, to its theses and dissertations, as well as to other kinds of grey literature*” (Pavlovska, 2018).

A survey at the University of Minnesota concludes that “*endorsing the depositing of materials into (institutional repositories) is a step in the direction of efficiency*”. Because many researchers are using Google and Google Scholar to locate grey literature, librarians should encourage the deposit grey literature into repositories in order to “*mitigate the web instability factor to which grey literature falls prey*” and to increase accessibility “*when Google crawls (institutional repositories)*” (Cooper et al., 2019). Based on many years of interdisciplinary research on grey marine literature, MacDonald et al. (2020) demonstrate how open practices can increase the impact of grey literature on political decision-making processes; yet, MacDonald and his colleagues also alert that this impact can be affected by (too) massive quantity, multiplicity of formats and content, and by (perceived) uncertain quality.

Similar results on the dissemination of grey literature via open repositories can be found in studies from Algeria (Babori & Aknouche, 2020), France (Schöpfel et al., 2019, 2020 and 2021), the United States (Lipinski & Kritikos, 2018; Mack, 2020), and from the BRICS countries (Bansal, 2020). Our own research in France shows that grey literature represents about 30% of the national repository’s content. While Mack (2020) provides a model for the development of a campus-based open access program with a special focus on grey literature, Babori & Aknouche (2020) make recommendations how to improve the deposit, visibility, and usage of grey literature in institutional repositories. Lipinski & Kritikos (2018) assess the open access policies of 22 US iSchools, show how these policies can affect access to grey literature, and recommend best practices for grey literature in institutional repositories.

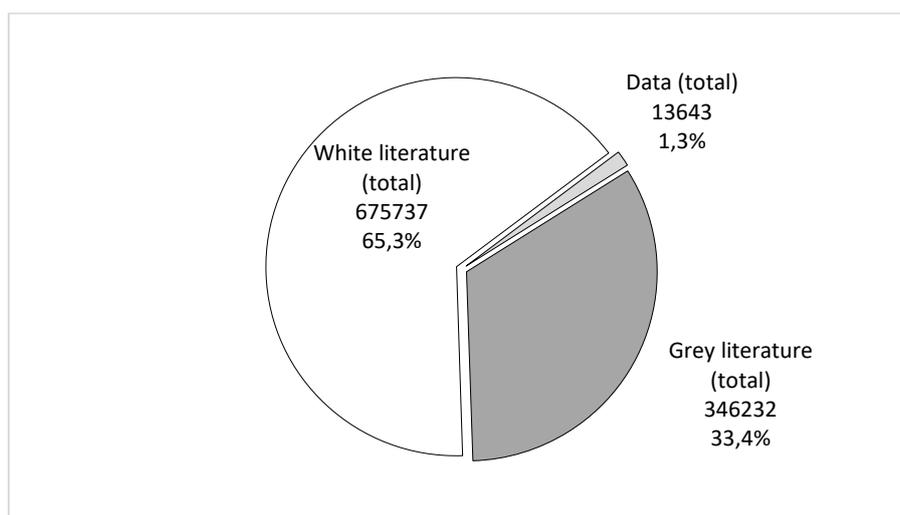


Figure 4. Part of grey literature in the HAL repository (source: Schöpfel et al., 2021)

Other papers describe the role of academic and national libraries (Richardson & Renner, 2018; Kingsley, 2020; Vandepontseele & Isbergue, 2020; Marsolek et al., 2021), with special attention on the need for identification, capture and deposit of grey literature into institutional repositories, for metadata curation, and for advocacy and awareness raising among the scientific communities. A comparison between institutional repositories and national aggregator systems insists on the need for advocacy and promotion on the institutional level: without the researchers’ practice of self-archiving, the repositories will remain empty (Cernohlavkova & Vycitalova, 2018).

The main argument of Marsolek and her colleagues from the University of Minnesota is that grey literature increases the richness and the diversity of the libraries digital collections which is crucial for the quality of systematic reviews. Also, they conclude with a kind of pro-grey pleading: *“Librarians have done and continue to do the important work of collecting GL, preserving it, and providing access to it. For those who are collecting GL, in solidarity we thank you and ask you to please continue the work (...) We encourage those who are not collecting GL to begin talking with your colleagues about how collecting it could be an opportunity for your organizations (...)”* (Marsolek et al., 2021).

A final, general comment, based on two very different papers. The first paper, an introduction into the field and concept of grey literature, emphasizes its importance and dynamic nature and concludes that *“technology has played a catalytic role in eradicating the issues and problems (of) grey literature”* (Gul et al., 2020), where technology means above all repositories (platforms) and search engines. A second paper presents results of a needs assessment for improving library support for dentistry researchers: 80% of participants use grey literature for their research (primarily, conference and seminar proceedings, and theses and dissertations), none of them consider it as “very important”, and only 20% indicate that they found it somewhat easy to search for and access it. Also, the authors of the study conclude that more and better library support is needed (He et al., 2018).

Taken together, these recent papers on grey literature in open repositories seem to confirm what has already been observed more than ten years ago (Schöpfel et al., 2012): repositories are essential for the dissemination, visibility, and use of grey literature, and their technology has contributed to expand its impact; but repositories have not (with the words of Gul et al., 2020) *“eradicating the issues and problems of grey literature”*, and they have not “white-washed” it. Open access or even accessibility is still not enough. More efforts and investment are required, to improve the curation of grey literature (metadata, including persistent identifiers) and to increase the assurance regarding quality and long-term archiving. The situation has not fundamentally changed since then; even if today, these questions are addressed (and reformulated) in terms of the FAIR principles (see below). One interesting example how the situation can be improved on the top of open repositories is presented by Vicary & Pettman (2020), capitalizing *“on the opportunities presented by Open Science and new technologies”*.

### **The transformation of research assessment (evaluation)**

How do the recent initiatives for a new system of research assessment affect the grey literature? Do they? Neither the San Francisco Declaration on Research Assessment (DORA)<sup>5</sup>, nor the recent European Commission Scoping report *Towards a reform of the research assessment system*<sup>6</sup>, nor the UNESCO *Recommendation for Open Science*<sup>7</sup>, nor the *Paris Call on Research Assessment* presented at the *Open Science European Conference* (OSEC)<sup>8</sup>, organized in the context of the French Presidency of the Council of the European Union, specifically mention grey literature. Does this mean that grey literature plays no (or only a marginal) role when it comes to transform the actual system dominated by journal-based metrics like the impact factor?

The reverse is true. All these declarations and reports have in common that they recommend diversity and inclusivity as fundamental criteria for research assessment. DORA (which has been signed by more than 22,000 institutions and individuals in 159 countries and is supported by the European Commission) recommends that funders and institutions consider, for the purposes of research assessment, *“the value and impact of*

<sup>5</sup> DORA (2012) <https://sfedora.org/>

<sup>6</sup> EC Scoping report (2021) <https://op.europa.eu/en/publication-detail/-/publication/36ebb96c-50c5-11ec-91ac-01aa75ed71a1/language-en>

<sup>7</sup> UNESCO Recommendation (2021) <https://en.unesco.org/science-sustainable-future/open-science/recommendation>

<sup>8</sup> OSEC (2022) <https://osec2022.eu/paris-call/>

*all research outputs (including datasets and software) in addition to research publications*". The EC Scoping report asks that research assessment should acknowledge the *"diversity of research activities and practices, with a diversity of outputs"*, while the Paris Call promotes the concept of bibliodiversity (see above), which means to consider *"all scientific productions"*, including for instance, preprints.

The UNESCO paper is more precise; starting from the assumption that *"evaluation systems should take into account the wide breadth of missions within the knowledge creation environment (and that) these missions come with different forms of knowledge creation and communication, not limited to publishing in peer reviewed international journals"*, the UNESCO recommends that research assessment should include all kind of scientific publications such as, among others, peer-reviewed journal articles and books, research reports and conference papers, disseminated by publishers and/or deposited in open repositories.

Our review retrieved only few papers that explicitly deal with this topic. The most relevant one is perhaps the case study by Giannini et al. (2017)<sup>9</sup> who analyzed the Italian research assessment exercises 2004-2010 and 2011-2014 and concluded that

- Both exercises did not explicitly include or mention grey literature;
- Yet, a detailed and time-consuming assessment allows the identification of some types of grey literature among the cited references of research products (output);
- This identified grey literature represents less than 1% of all cited output;
- This percentage is slightly increasing, from 0.61% (2004-2010) to 0.74% (2010-2014);
- The relative importance of grey literature varies between disciplines.

These findings are not optimistic but appear realistic, and they are compliant with the general observation that grey literature is *"usually ignored in research evaluations"* (Bickley et al., 2020). However, without any other empirical evidence (we could not retrieve any other paper like this one), it is impossible to say if these findings are representative for the research assessment in large research organizations, and if there have been changes since then.

A recent paper by Reed et al. (2021) proposes a conceptual framework of research assessment exercises. Based on a systematic review, the authors identify five different approaches of research impact evaluation; but only one design explicitly includes grey literature, i.e., *"evidence synthesis approaches"* which provide summative evidence of research as a sufficient cause of impact (figure 5).

---

<sup>9</sup> This paper was retrieved because the communication was published in the conference proceedings in 2018.

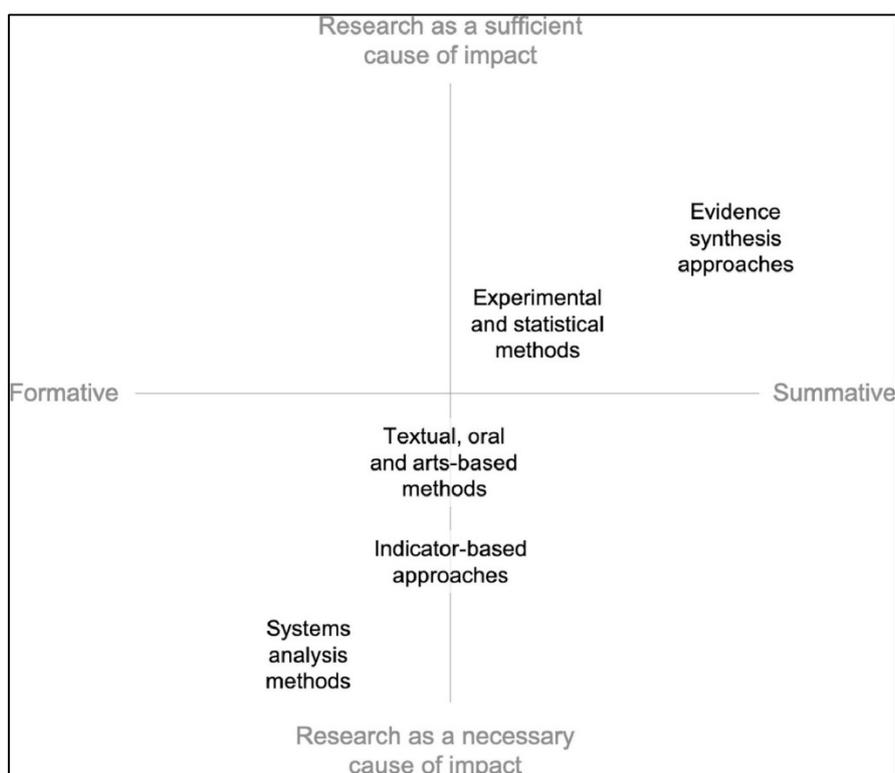


Figure 5. Five types of impact evaluation designs (Source: Reed et al., 2021)

Following Reed et al. (2021), such evidence synthesis “typically takes place at the programme level and draws on bodies of work emerging from multiple projects”; it is “a process of carrying out a review of existing data, literature and other forms of evidence with pre-defined methodological approaches, to provide a transparent, rigorous and objective assessment of whether something arising from research is a sufficient cause of impactful outcomes (...)”. As systematic reviews generally include grey literature, they are (or should be) part of this kind of research assessment. However, the authors are aware of the potential issues of including grey literature into evidence synthesis: “it is time and labour intensive as it requires considerable consultation with likely end-users and searching of unpublished and grey literature, often by hand and often at geographically disparate locations”.

If the interest of grey literature for systematic reviews is largely accepted (Bonato, 2018; Hoffecker, 2020; Schöpfel & Prost, 2021), the related problems have been identified, like the lack of persistent identifiers and of rich metadata, reduced findability and accessibility, issues that limit the grey literature’s usefulness for monitoring and altmetrics and make it still more difficult to conduct scientometric assessment with grey literature than with journal articles or books (Bickley et al., 2020; Roos Lindgreen et al., 2021; Schöpfel & Prost, 2020; Valles et al., 2020). Also, when Kenfield et al. (2019) tried to develop a framework for the reuse assessment of grey literature in institutional repositories, they observed different barriers, such as “lack of best practices, documented workflows, assessment training, and staffing”. “Improper indexing and bibliographic control” are other barriers to the reuse (citation) and assessment of grey literature (Shrivastava & Mahajan, 2020).

Obviously, we can observe a kind of paradox: while the international initiatives to transform quantitative and journal-based research assessment explicitly call for more diversity of research outputs, the role and importance of grey literature for research assessment remain marginal, so far. Is this definitive and inherent to the characteristics of grey literature? Not sure. But without any doubt, advocacy and promotion of diversity will not be enough to raise awareness for the interest of grey literature for research

assessment, if it is not (more) compliant with the fundamental values and principles of open science, including FAIR. The UNESCO recommendation is very clear: research assessment should make use of repositories and other platforms “*that are supported and maintained by an academic institution, scholarly society, government agency or other well-established not-for-profit organization devoted to common good that enables open access, unrestricted distribution, interoperability and long-term digital preservation and archiving*”; and it should ensure diversity in “*scholarly communications with adherence to the principles of open, transparent and equitable access and supporting non-commercial publishing models and collaborative publishing models with no article processing charges or book processing charges*”. To increase its role for research assessment, grey literature must be open, accessible, and findable, and it must be disseminated on trustworthy platforms and with FAIR data and metadata.

### **The application of FAIR principles (processing)**

Are the FAIR principles relevant for the processing of grey literature, and if so, which ones and in which way? Initially developed for research data repositories (Wilkinson et al., 2016), they have been progressively applied to all kind of infrastructures, procedures, and resources, including items belonging to grey literature<sup>10</sup>. Our systematic review reveals that above all, two FAIR principles are an issue for grey literature, i.e., findability and accessibility.

**Findability:** Grey literature is relevant if not a must-have for systematic reviews (see for instance Whaley et al., 2020; also, Schöpfel & Prost, 2021); yet, the inclusion of grey literature in reviews is generally considered as more time-consuming and costly as for journal articles (Enticott et al., 2018), mainly because grey resources have less standardized publication formats (Landerdahl Stridsberg et al., 2022). A couple of papers make recommendations how to improve findability of grey literature, especially through the enrichment of metadata and indexing (Jamouille et al., 2017; Marsolek et al., 2018) and the minting and usage of persistent identifiers, namely DOI but also ISSN, ORCID and ROR (Farace et al., 2019 and 2021; Price & Murtagh, 2020; Reynolds & Ross, 2020). Careful curation is required to improve the visibility and searchability of grey literature; or with the words of Marsolek et al. (2018), an “*increased level of collection development of grey literature in institutional repositories would facilitate preservation and increase the findability and reach of grey literature*”.

**Accessibility:** Open access with no or less restrictions remains an issue for grey literature, even if the deposit of theses and dissertations, conference proceedings, working papers and reports on open repositories improved their dissemination and impact. As said above, many repositories contain grey literature (figure 1) where it represents up to 30% of the content. However, other platforms are useful for different types of grey literature, such as PubMed Central and Clinicaltrials.gov from the US National Library of Medicine (Sheehan, 2021) or ConfIDent, the conference platform from the German National Library for Scientific and Technical Information (Hagemann-Wilholt et al., 2019). On the other hand, embargo periods can reduce the accessibility of grey resources on institutional repositories (Rasuli et al., 2022).

---

<sup>10</sup> See for instance the recent NDLT conference on the FAIRness of electronic theses and dissertations at Novi Sad, Serbia <https://etd2022.uns.ac.rs/>

### Who has an embargo policy?

Yes: 43  
No: 57

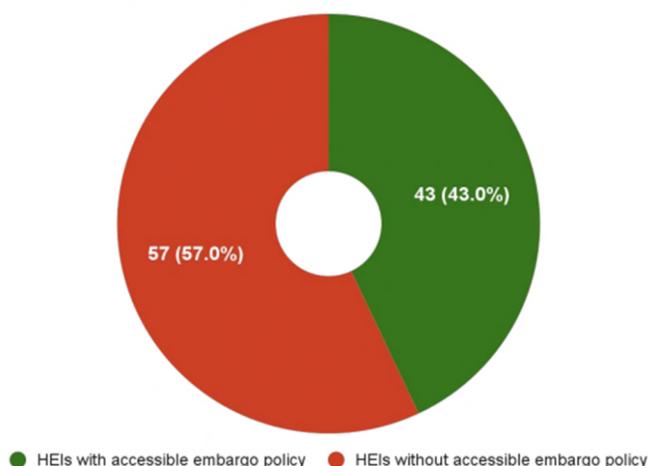


Figure 6. Embargo policies in universities (source: Rasuli et al., 2022)

We can make two recommendations, and they are not really new: The producers of academic grey literature (research organizations, universities...) and repository managers should systematically attribute persistent identifiers to grey literature (dissertations, reports, working papers, conference papers...), especially when these documents are deposited in institutional repositories. On the other hand, institutions and repositories should create a framework (= technical, political, social, and legal environment) that fosters openness and accessibility of grey literature

### Perspectives

The findings of our review reflect a continuing scientific and professional interest in grey literature, in- and outside of the GreyNet community of practice, as reflected by special issues in academic journals (Dyas-Correia & Mering, 2021; Garousi et al., 2022), by handbooks (Bonato, 2018) and by a steadily growing number of systematic reviews (Schöpfel & Prost, 2021). The main issue of grey literature today is not typology but FAIRness.

Open science improves the openness of grey literature, its visibility, and its accessibility especially in open repositories. However, as our findings show, open science also emphasizes and exacerbates the usual shortfalls of grey literature, the lack of persistent identifiers, problems with the quality and the richness of metadata, and with the long-term conservation. According to some recent research, these issues are not specific to grey literature; inconsistent implementation of unique identifiers are key challenges for open access books (Laakso, 2022), and open access journals can simply vanish from the web (Laakso et al., 2021). But the lack of long-term conservation, of findability (i.e., non-efficient search and retrieval) and of standards are serious threats to the sustainability of grey literature (Savic, 2022).

Thus, open science is an opportunity as well as a challenge for grey literature. One way to face the challenge is the continuous work on the concept of grey literature: *“Taking into consideration the volume and speed of (grey literature) creation, there seems to be a need to revisit the old definition of GL by refocusing on quality, intellectual property, curation, sustainability and usability (so that) proper attention can be focused on relevant (...) issues and solutions”* (Savic, 2018). This conceptual work is not simple, as there are contradictory arguments and perspectives; but as open science is here to stay, grey literature should be *“completely integrated in the emerging open science infrastructures, as open as possible (and just as closed as necessary), easily retrievable and accessible and largely reusable by (machines)”* (Schöpfel & Rasuli, 2018).

In the previous sections we made some recommendations for grey literature producers and services. Our final suggestion for the GreyNet community of practice would be a kind of grey literature observatory, a virtual place for information professionals, researchers and service providers, to analyze and discuss the impact of open science on grey literature and to contribute to its integration into this new landscape, with advocacy, information, and training.

## Bibliography

- Artus, H. (2003). Old WWine in New Bottles? Developments in Electronic Information and Communication: Structural Change and Functional Inertia. *Fifth International Conference on Grey Literature*. December 4-5, 2003, Amsterdam.
- Babori, A., & Aknouche, N. (2020). Increasing the Visibility of Grey Literature in Algerian Institutional Repositories. *The Grey Journal*, 16(Special Issue), 43-52. <http://dspace.univ-constantine2.dz/handle/123456789/3>
- Bansal, S. (2020). Grey Literature Archiving Pattern in BRICS Open Access Repositories. *The Grey Journal*, 16(3), 199-203.
- Bonato, S. (2018). *Searching the grey literature: a handbook for finding reports, working papers, and unpublished research*. Lanham, Maryland: Rowman & Littlefield.
- Bickley, M. S., Kousha, K., & Thelwall, M. (2020). Can the impact of grey literature be assessed? An investigation of UK government publications cited by articles and books. *Scientometrics*, 125(2), 1425–1444. <https://doi.org/10.1007/s11192-020-03628-w>
- Bielskas, A. S., Holmes, E., Jackson, E. M., Wacker, M., & Mercurio, J. R. (2022). Technical Reports in the Repository: Syncing a Unique Grey Literature Collection in Multiple Columbia University Library Systems. *GSIS Newsletter*, 293 (April/May), 11-16. <https://doi.org/10.7916/dg8c-hd03>
- Černohlávková, P., & Vyčítalová, H. (2018). Collecting grey literature – Institutional repository versus national aggregator. *The Grey Journal*, 14(2), 81–86.
- Cirkovic, S., Schöpfel, J., & Prost, H. (2019). What about ODTs? Are They Grey? Some Comments on Scientific Heritage and Grey Literature. *The Grey Journal*, 15(Special Issue), 27-35.
- Cooper, K., Marsolek, W., Riegelman, A., Farrell, S., & Kelly, J. (2019). Grey Literature: Use, Creation, and Citation Habits of Faculty Researchers across Disciplines. *Journal of Librarianship and Scholarly Communication*, 7(1). <https://doi.org/10.7710/2162-3309.2314>
- Dyas-Correia, S., & Mering, M. (2020). Introduction. *The Serials Librarian*, 79(3–4), 229–230. <https://doi.org/10.1080/0361526X.2020.1866948>
- Enticott, J., Buck, K., & Shawyer, F. (2018). Finding ‘Hard to Find’ Literature on Hard to Find Groups: A Novel Technique to Search Grey Literature on Refugees and Asylum Seekers. *International Journal of Methods in Psychiatric Research*, 27(1). <https://doi.org/10.1002/mpr.1580>
- Farace, D., Frantzen, J., Biagioni, S., Carlesi, C., De Robbio, A., Češarek, A., ... Crowe, J. (2019). AccessGrey: Securing Open Access to Grey Literature for Science and Society. *Twentieth-First International Conference on Grey Literature*. October 22-23, 2019, Hannover. <https://doi.org/10.26069/GREYNET-2020-000.219-GG>
- Farace, D., Biagioni, S., Carlesi, C., & Baars, C. (2021). Grey Literature and Persistent Identifiers: GreyNet’s Use Case. *Twenty-Third International Conference on Grey Literature*. December 6-7, 2021, Amsterdam. <https://doi.org/10.5446/55728>
- Garousi, V., Rainer, A., Felderer, M., & Mäntylä, M. V. (2022). Introduction to the Special Issue on: Grey Literature and Multivocal Literature Reviews (MLRs) in software engineering. *Information and Software Technology*, 141, 106697. <https://doi.org/10.1016/j.infsof.2021.106697>
- Gelfand, J. M., & Lin, A. (2020). How open science influences next developments in grey literature. *The Grey Journal*, 16(1), 34–48.
- Giannini, S., Deluca, R., Molino, A., & Biagioni, S. (2017). Grey Literature and Research Assessment exercises: From the current criteria to the Open Science models. *Nineteenth International Conference on Grey Literature*. October 23-24, 2017, Rome. <https://doi.org/10.26069/GREYNET-2018-000.001-GG>
- Goggi, S., Pardelli, G., Russo, I., Bartolini, R., & Monachini, M. (2017). Providing Access to Grey Literature: The CLARIN Infrastructure (Poster). *Nineteenth International Conference on Grey Literature*. October 23-24, 2017, Rome.
- Gul, S., Shah, T. A., Ahmad, S., Gulzar, F., & Shabir, T. (2020). Is Grey Literature Really Grey or a Hidden Glory to Showcase the Sleeping Beauty. *Collection and Curation*, 40(3), 100–111. <https://doi.org/10.1108/CC-10-2019-0036>
- Hagemann-Wilholt, S., Hauschke, C., & Plank, M. (2019). ConfIDent – An Open Platform for FAIR Conference Metadata. *Twentieth-First International Conference on Grey Literature*. October 22-23, 2019, Hannover. <https://doi.org/10.26069/greynet-2020-000.216-GG>

- He, H. Y., Gerbig, M., & Kirby, S. (2019). Needs assessment for improving library support for dentistry researchers. *Journal of the Medical Library Association: JMLA*, 107(3), 352–363. <https://doi.org/10.5195/jmla.2019.556>
- Hoffecker, L. (2020). Grey Literature Searching for Systematic Reviews in the Health Sciences. *The Serials Librarian*, 79(3–4), 252–260. <https://doi.org/10.1080/0361526X.2020.1847745>
- Jamouille, M., Cardillo, E., Ittoo, A., Vander Stichele, R., Resnick, M. P., Grosjean, J., & Vanmeerbeek, M. (2017). Indexing grey literature in General Practice: Family Medicine in the Era of Semantic Web. *Nineteenth International Conference on Grey Literature*. October 23-24, 2017, Rome. <http://greyguide.isti.cnr.it/index.php/34-gl-slide-show/gl19-slide-show/59-gl19-slide>
- Kenfield, A. S., Kelly, E., Muglia, C., O’Gara, G., Thompson, S., & Woolcott, L. (2019). Measuring reuse of institutionally-hosted grey literature. *The Grey Journal*, 15(1), 51–58.
- Kingsley, D. (2020). The ‘Impact Opportunity’ for Academic Libraries through Grey Literature. *The Serials Librarian*, 79(3–4), 281–289. <https://doi.org/10.1080/0361526X.2020.1847744>
- Kousha, K., Thelwall, M., & Bickley, M. (2022). The High Scholarly Value of Grey Literature Before and During Covid-19. *Scientometrics*, 127(6), 3489–3504. <https://doi.org/10.1007/s11192-022-04398-3>
- Kumazaki, Y., Suzuki, S., Kanazawa, M., Kunii, K., Yonezawa, M., & Itabashi, K. (2019). Published Electronic Media are Becoming Grey. *The Grey Journal*, 15(1), 45-49.
- Laakso, M. (2022). *Open access books through open data sources: Assessing prevalence, providers, and preservation*. Preprint. <https://doi.org/10.5281/zenodo.7305490>
- Laakso, M., Matthias, L., & Jahn, N. (2021). Open is not Forever: A Study of Vanished Open Access Journals. *Journal of the Association for Information Science and Technology*, 72(9), 1099–1112. <https://doi.org/10.1002/asi.24460>
- Landerdahl Stridsberg, S., Richardson, M. X., Redekop, K., Ehn, M., & Wamala Andersson, S. (2022). Gray Literature in Evaluating Effectiveness in Digital Health and Health and Welfare Technology: A Source Worth Considering. *Journal of Medical Internet Research*, 24(3), e29307. <https://doi.org/10.2196/29307>
- Lipinski, T. A., & Kritikos, K. C. (2018). How Open Access Policies Affect Access to Grey Literature in University Digital Repositories: A Case Study of iSchools. *The Grey Journal*, 14(1), 6–20.
- Lombardi, S. (2021). Every Document is Born ‘Grey’ - Some Documents Can Become ‘Open’. *The Grey Journal*, 17(2), 77-80.
- Luzi, D. (2010). Grey Documents in Open Archives. *The Grey Journal*, 6(3), 137-144.
- MacDonald, B. H., Cadman, R., Martin, C., Ryder-Burbidge, S., Soomai, S. S., Stewart, I., & Wells, P. G. (2020). Is the Production and Use of Grey Marine Literature a Model for Open Science? *The Grey Journal*, 16(2), 73-83.
- Mack, D. C. (2020). Open Access in the Academy: Developing a Library Program for Campus Engagement. *The Grey Journal*, 16(3), 181–185.
- Mackenzie Owen, J. (1997). Expanding the Horizon of Grey Literature. *Third International Conference on Grey Literature*. November 13-14, 1997, Luxemburg. <http://eprints.rclis.org/5654/>
- Marsolek, W. R., Cooper, K., Farrell, S. L., & Kelly, J. A. (2018). The Types, Frequencies, and Findability of Disciplinary Grey Literature within Prominent Subject Databases and Academic Institutional Repositories. *Journal of Librarianship and Scholarly Communication*, 6(1). <https://doi.org/10.7710/2162-3309.2200>
- Marsolek, W., Farrell, S. L., Kelly, J. A., & Cooper, K. (2021). Grey Literature: Advocating for Diverse Voices, Increased Use, Improved Access, and Preservation. *College & Research Libraries News*, 82(2), 58. <https://doi.org/10.5860/crln.82.2.58>
- Pavlovská, E. (2018). “Grey Literature” in Electronic Archives. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 8, 105–108. <https://doi.org/10.55630/dipp.2018.8.8>
- Price, R., & Murtagh, J. (2020). An Institutional Repository Publishing Model for Imperial College London Grey Literature. *The Serials Librarian*, 79(3–4), 349–358. <https://doi.org/10.1080/0361526X.2020.1847737>
- Puccinelli, R., Reggiani, L., Saccone, M., & Truffelli, L. (2020). Open Educational Resources and Library & Information Science: Towards a Common Framework for Methodological Approaches and Technical Solutions. *The Grey Journal*, 16(3), 161–171.
- Rasuli, B., Schöpfel, J., Boock, M., & van Wyk, B. (2022). Reasons for Embargoing Academic Publications in Institutional Repositories. Case Study of Theses and Dissertations. *OR 2022. The 17th International Conference on Open Repositories*. June 6-9, 2022, Denver, CO. <https://hal.univ-lille.fr/hal-03758331>
- Reed, M. S., Ferré, M., Martín-Ortega, J., Blanche, R., Lawford-Rolfe, R., Dallimer, M., & Holden, J. (2021). Evaluating impact from research: A methodological framework. *Research Policy*, 50(4), 104147. <https://doi.org/10.1016/j.respol.2020.104147>
- Reynolds, R. R., & Ross, K. E. A. (2020). ISSN Is for Black, White, and Many Shades of Grey. *The Serials Librarian*, 79(3–4), 290–298. <https://doi.org/10.1080/0361526X.2020.1847746>
- Richardson, L. M., & Renner, B. R. (2018). Capturing Gray Literature for the Institutional Repository. *Sixth Entrepreneurial Librarian Conference*. October 12, 2018, Wake Forest University, Winston-Salem, NC. <https://doi.org/10.17615/wm79-jv24>

- Roos Lindgreen, E., Mondello, G., Salomone, R., Lanuzza, F., & Saija, G. (2021). Exploring the effectiveness of grey literature indicators and life cycle assessment in assessing circular economy at the micro level: a comparative analysis. *The International Journal of Life Cycle Assessment*, 26(11), 2171–2191. <https://doi.org/10.1007/s11367-021-01972-4>
- Savic, D. (2018). Rethinking the Role of Grey Literature in the Fourth Industrial Revolution. *The Grey Journal*, 14(2018), 7–14.
- Savic, D. (2022). The Impact of Digital Transformation on the Sustainability of Grey Literature. *The Grey Journal*, 18(1), 7-11.
- Schöpfel, J. (2006). Observations on the Future of Grey Literature. *The Grey Journal*, 2(2), 67–76. [https://hal.archives-ouvertes.fr/sic\\_00168998](https://hal.archives-ouvertes.fr/sic_00168998)
- Schöpfel, J. (2019). Grey Literature and Professional Knowledge Making. In L. Börjesson & I. Huvila (eds.), *Research Outside the Academy: Professional Knowledge-Making in the Digital Age* (pp. 137–153). Cham, Palgrave Macmillan. [https://doi.org/10.1007/978-3-319-94177-6\\_8](https://doi.org/10.1007/978-3-319-94177-6_8)
- Schöpfel, J., Le Bescond, I., & Prost, H. (2012). Open is Not Enough: A Case Study on Grey Literature in an OAI Environment. *The Grey Journal*, 8(2), 112-124.
- Schöpfel, J., & Farace, D. (2018). Grey Literature. In J. D. McDonald & M. Levine-Clark (Eds.), *ELIS Encyclopedia of Library and Information Sciences* (4th edition). Boca Raton FL, CRC Press. <https://doi.org/https://doi.org/10.1081/E-ELIS4>
- Schöpfel, J., & Rasuli, B. (2018). Are Electronic Theses and Dissertations (Still) Grey Literature in the Digital Age? A FAIR Debate. *The Electronic Library*, 36(2), 208–219. <https://doi.org/10.1108/EL-02-2017-0039>
- Schöpfel, J., Prost, H., & Ndiaye, E.I. (2019). Going Green. Publishing Academic Grey Literature in Laboratory Collections on HAL. *Twentieth-First International Conference on Grey Literature*. October 22-23, 2019, Hannover. <https://hal.archives-ouvertes.fr/hal-02300017>
- Schöpfel, J., Kergosien, E., Prost, H., & Thiault, F. (2020). The Grey Side of the Green Road. Empirical Assessment of Academic Publishing in the HAL Open Repository. *Twentieth-Second International Conference on Grey Literature*. November 19, 2020, Rome. <https://doi.org/10.26069/greynet-2021-000.460-gg>
- Schöpfel, J., & Prost, H. (2020). The scope of open science monitoring and grey literature. *The Grey Journal*, 16(Special Issue Winter), 62–71.
- Schöpfel, J., & Prost, H. (2021). How Scientific Papers Mention Grey Literature: A Scientometric Study Based on Scopus Data. *Collection and Curation*, 40(3), 77–82. <https://doi.org/10.1108/CC-12-2019-0044>
- Schöpfel, J., Kergosien, E., Prost, H., & Thiault, F. (2021). Grey Literature in Open Repositories: New Insights and New Issues. *Twenty-Third International Conference on Grey Literature*. December 6-7, 2021, Amsterdam. <https://doi.org/10.5446/55709>
- Schöpfel, J., Farace, D., Baxter, D., Giannini, S., Molino, A., Lipinski, T., ... Savic, D. (2022). Data from “Exploring Next Generation Grey.” *The Grey Journal*, 18(1), 55–62. <https://doi.org/10.17026/dans-xrg-2gfg>
- Sheehan, J. (2021). Open Science and the Transformation of Scholarly Communication. *Twentieth-Second International Conference on Grey Literature*. November 19, 2020, Rome. <https://doi.org/10.26069/greynet-2021-000.454-GG>
- Shrivastava, R., & Mahajan, P. (2020). Analysis of the usage and diversity of grey literature in addiction research: a study. *Collection and Curation*, 40(3), 93–99. <https://doi.org/10.1108/CC-12-2019-0046>
- Valles, M., Injante, R., Hernández, E., Riascos, J., Galvez, M., & Velasco, J. (2020). An Altmetric Alternative for Measuring the Impact of University Institutional Repositories’ Grey Literature. In: R. Mugnaini (Ed.), *Data and Information in Online Environments*. DIONE 2020. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 319. Cham, Springer (222–234). [https://doi.org/10.1007/978-3-030-50072-6\\_17](https://doi.org/10.1007/978-3-030-50072-6_17)
- Vandepontseele, S., & Isbergue, N. (2020). Grey Literature and Legal Deposit: The Approach of the Royal Library of Belgium. *The Serials Librarian*, 79(3–4), 299–306. <https://doi.org/10.1080/0361526X.2020.1847742>
- Vicary, T., & Pettman, I. (2020). Abstracting and Indexing as an Enabling Interface between Open Science and Grey Literature – The Approach of the Aquatic Sciences and Fisheries Abstracts Service. *The Grey Journal*, 16(2), 84–93.
- Whaley, P., Aiassa, E., Beausoleil, C., Beronius, A., Bilotta, G., Boobis, A., ... Halsall, C. (2020). Recommendations for the Conduct of Systematic Reviews in Toxicology and Environmental Health Research (COSTER). *Environment International*, 143, 105926. <https://doi.org/10.1016/j.envint.2020.105926>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>
- Williams, C. C. (2020). Discovering and Identifying Grey Literature in the Field of Law. *The Serials Librarian*, 79(3–4), 245–251. <https://doi.org/10.1080/0361526X.2020.1847741>

**Appendix**

We searched for publications in English, published between 2018 and 2022.

**Web of Science:** We searched in the WoS Core Collection for “grey literature” or “gray literature” in title, abstract and keywords. Review articles were excluded. We then filtered with the following terms in the full text: “open science”, “repositories”, “bibliodiversity”, “assessment”, “evaluation”, “FAIR principles” or “DORA”.

**Scopus:** We searched for “grey literature” or “gray literature” in title, abstract and keywords. Review articles were excluded. We then filtered with the following terms in the full text: “open science”, “repositories”, “bibliodiversity”, “assessment”, “evaluation”, “FAIR principles” or “DORA”.

**GreyGuide:** Manually selection of relevant conference papers.

**Dimensions:** We searched for “grey literature” or “gray literature” in title and abstract. We then filtered with the following terms in the full text: “open science”, “repositories”, “bibliodiversity”, “assessment”, “evaluation”, “FAIR principles” or “DORA”. We selected all documents in Library and Information Science.

**Google Scholar:** We searched for “grey literature” or “gray literature” in title. We excluded all citations. We then filtered with the following terms in the full text: “open science”, “repositories”, “bibliodiversity”, “assessment”, “evaluation”, “FAIR principles” or “DORA”..

## Data Science as a Research Support Service and the Role of the Libraries: UF's CNI Executive Roundtable Experience

Plato L. Smith, Erik Deumens, and Christopher Barnes

University of Florida; George A. Smathers Libraries, USA

### Video Presentation

<https://av.tib.eu/media/58295>

### Abstract

**Design:** The author used a stakeholder approach to address ten (10) questions set by the program sponsor, Coalition for Networked Information (CNI), for institutional participation in a CNI executive roundtable.

**Purpose:** The goal of the CNI Executive Roundtable was to gain insight into how data science expertise is being provided to the university community in the broader context of research support services at key CNI member institutions. This case study shares insights from a large R1 institution in the southeast United States of America.

**Methodology:** The author used a participatory action research method approach.

**Findings:** Participation in the CNI Executive Roundtable discussion revealed the need for interdisciplinary, socio-technical collaborations of key stakeholders within and across the university community and campus units to further develop, maintain, and sustain the developing data science ecosystem as a research support service and the role of libraries.

**Research implications:** Data science support requires collaborations across many units.

**Practical implications:** Collaborations across campus units support interdisciplinary research which is necessary for developing data science as a research support service.

**Originality/value:** The libraries collaborated with information technology research computing and clinical and translational science information technology units to answer roundtable questions which were shared with CNI after the March 16, 2022 meeting.

### Introduction

The Dean of the Libraries sometimes shares professional development opportunities received through list serves affiliated with key senior administrators' communities of practice. On February 14, 2022, the CNI-MEMBER-REPS list serve received a notification titled REMINDER: Data Science Research Support & Libraries CNI Exec. Roundtable. The Dean forwarded the notification to the Data Management Librarian for consideration of participation. After reviewing the notification, the Data Management Librarian informed the Dean of Libraries on February 15, 2022 that the Libraries should express an interest to participate in the roundtable.

The deadline to express interest in the CNI Executive Roundtable Data Science as a Research Support Service and the Role of Libraries was February 21, 2022. Any CNI member affiliate could apply to participate in the Roundtable either as one individual or a team of up to three individuals who have different roles (e.g. a library director, a CIO, a head of research computing, a chief research officer or a faculty officer or a faculty data science leader) to represent the institution. The team representing the author's institution included the Library Director, Director of Information Technology Research Computing, and the Data Management Librarian.

The express of interest application form included submission of a brief (2-3 sentences) of the institution's primary interest in the Roundtable and following questions: What is the scope of data science research support services at your institution? Within the library? How much demand is there for research support services and how are you managing it? An expression of interest to participate in the Roundtable was submitted on February 19,

2022. The request to participate was accepted on February 24, 2022. The author's team attended the Roundtable on March 16, 2022.

The author used a participatory action research approach. Participatory action research is "a type of research in which the researcher involves some organizational members as active participants throughout the process of studying an organization; the goal is making changes in the organization." (Schutt, 2006, p. 1-26).

"The Coalition for Networked Information (CNI) is dedicated to supporting the transformative promise of digital information technology for the advancement of scholarly communication and the enrichment of intellectual productivity. Over 200 institutions representing higher education, publishing, information technology, scholarly and professional organizations, foundations, and libraries and library organizations make up CNI's members; CNI is entirely funded through membership dues. Semi-annual membership meetings bring together representatives of CNI's constituencies to discuss ongoing and new projects and to plan for future initiatives<sup>1</sup>." The goal of the Executive Roundtable was to gain insight into how data science expertise is being provided to the university community in the broader context of research support services at key CNI member institutions. CNI is particularly interested in what roles, if any, the university library is playing in data science as a research support service. Each participating institutional team were informed to be prepared to speak to at least most of the research support questions. Out of the thirteen teams that participated, this article articulates responses from the author's institution. "The goal of the Executive Roundtable was to gain insight into how data science expertise is being provided to the university community in the broader context of research support services at key CNI member institutions. CNI was particularly interested in what roles, if any, the university library is playing in research support. To enable Roundtable participants to fully understand the context at each institution, participating institutional teams prepared to speak to at least most of the following issues<sup>2</sup>" covered in the ten (10) questions provided to the participating institutions in the roundtable discussion. The University of Florida was one of 13 key CNI member institutions selected to participate in this executive roundtable discussion.

### Literature Review

The National Institutes of Health Office of Data Science Strategy (<https://datascience.nih.gov>) includes several data science themes. The data science themes are data infrastructure, data ecosystem, tools and analytics, community engagement, and workforce development. The Committee on Data International Science Council (CODATA) includes a CODATA-RDA (Research Data Alliance) Schools of Research Data Science to address the need for foundational data science skills in all disciplines (<https://codata.org/initiatives/data-skills/research-data-science-summer-schools/>). "According to CODATA-RDA, research data science requires a range of skills relating to data in order to address some contemporary research challenges. "Contemporary research – particularly when addressing the most significant, inter-disciplinary research challenges – cannot effectively be done without a range of skills relating to data, including the principles and practice of Open Science and research data management and curation, the use of a range of data platforms and infrastructures, large scale analysis, statistics, visualization and modelling techniques, software development and annotation, etc., etc." (CODATA, 2022). Thus, data science as a research support service and the role of the libraries will require broad collaborations with stakeholders to develop capacity, infrastructures, and resources.

---

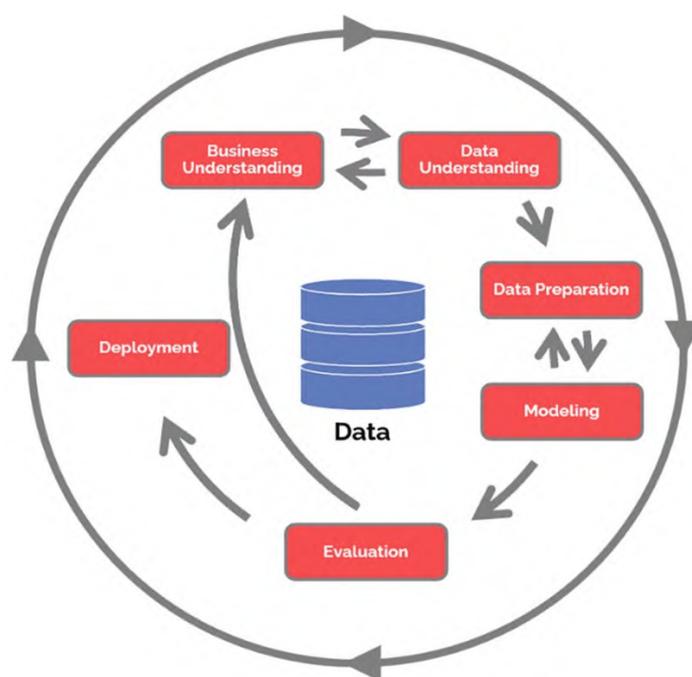
<sup>1</sup> About CNI <https://www.cni.org/about-cn>.

<sup>2</sup> CNI Executive Roundtable Call for Expressions of Interest, March 16, 2022, 2:00-4:30pm ET.

### What is data science?

Data science can be described as “a set of fundamental principles that support and guide the principled extraction of information and knowledge from data” (Provost and Fawcett, 2013). The **C**Ross **I**ndustry **S**tandard **P**rocess for **D**ata **M**ining (CRISP-DM)<sup>3</sup> “enable organizations to learn in concert with data (Shafique and Qaiser, 2014; Sharma, Osei-Bryson and Kasper, 2012)” (van der Voort et al., 2021, 1). The CRISP-DM “is a process model that serves as the base for a data science process” (Hotz, 2022). The CRISP-DM Life Cycle by Hotz (Figure 1) illustrates key data science phases. The key six sequential phases include:

1. Business understanding – What does the business need?
2. Data understanding – What do we have / need? Is it clean?
3. Data preparation – How do we organize the data for modelling?
4. Modelling – What modelling techniques should we apply?
5. Evaluation – Which model best meets the business objectives?
6. Deployment – How do stakeholders access the results?



CRISP-DM Diagram. Inspired by Wikipedia

Figure 1. CRoss Industry Standard Process for Data Mining (CRISP-DM) Diagram (Hotz, 2022)

### Developing research data science as a support services

The National Institutes of Health National Library of Medicine Network of the National Library of Medicine (NNLM) priority areas include: Citizen Science and Crowdsourcing, Data Science, Student Engagement, and Substance Use Disorders<sup>4</sup>. The NNLM data science priority area supports the goals of the National Library of Medicine Strategic Plan 2017-2027: A Platform for Biomedical Discovery and Data-Powered Health<sup>5</sup>. The three goals of strategic plan include: accelerate discovery and advance health through data-driven research, reach more people in more ways through enhanced dissemination and engagement, and build a workforce for data-driven research and health. Developing data science as a research support services and the roles of the libraries as collaborators, resource brokers, and stakeholders at Coalition for Networked Information (CNI) member institutions can support the NLM in meeting the evolving needs of the research and clinical communities via institutional support for resources, software, and tools. For

<sup>3</sup> Hotz, N. (2022). What is CRISP DM? <https://www.datascience-pm.com/crisp-dm-2/>.

<sup>4</sup> NNLM. (2022). Network of the National Library of Medicine Initiatives. <https://nnlm.gov/initiatives>.

<sup>5</sup> NLM. (2019). [https://www.nlm.nih.gov/pubs/plan/lrp17/NLM\\_StrategicReport2017\\_2027.html](https://www.nlm.nih.gov/pubs/plan/lrp17/NLM_StrategicReport2017_2027.html).

example, Emerald Cloud Lab (ECL), developed by two Carnegie Mellon University alumni, is “the world’s first Academic Cloud that allows you to remotely control a complete life sciences laboratory from your compute.”<sup>6</sup> ECL was referenced during keynote address by Clifford Lynch on Program Day 1 of the 42<sup>nd</sup> International Association of University Libraries (IATUL) on June 13, 2022. Electronic research notebooks (ERNs) also known as electronic lab notebooks (ELNs) are emerging as potential core resources for developing research data science as support services.

### **The emerging role of ELN/ERN in RDM as part of data science support**

On June 16, 2022, an international Research Notebooks Community Discussion meeting led by [University of Glasgow](#) and [Jisc](#) representatives introduced four case studies. The use cases were from the University of Oxford, Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), University College London, and The Clausthal University of Technology. Despite the differences in approaches to exploring electronic lab notebook (ELN) at the four institutions along with their current level of developments, all articulated the fact that ELNs are becoming important in the development of research support services for their researchers with one institution articulating there are “rumours that some of funding councils may ask for ELNs to be published as the end of research grant.” ELN<sup>7</sup> are also known as electronic research notebooks (ERNs) to denote use beyond traditional wet labs/lab environments.

The University College London (UCL) Research Data Lifecycle (see Figure 2) include eleven data management components which also represent the research data management and curation foundational skills that contribute to research data science. The eleven components illustrated in the UCL research data lifecycle are: Idea, Planning, Funding bid, Data collection, Data storage, Data analysis, Data deposit, Data production, Preservation & Curation, Discovery, and Access. Figure 1 describes the key processes in the research data lifecycle at UCL aligned with maturity of services corresponding to each research data lifecycle activity. The color-coded legend describes the level of maturity from green representing Mature on scale to red representing Immature. The greenish-yellow arrow represents the maturity of the DMPTool support service at the University of Florida (UF) whereas the red arrow represents the immaturity of ERNs support service at UF.

Figure 3 represents an ERN, RSpace (<https://www.researchspace.com>), mapped to UCL’s Research Data Cycle directly addressing the Planning, Data collection, Data Storage, Data deposit, Article/Data Publication, Preservation & Curation, Discovery, and Access activities. Figure 3 represents the data preparation, active research, and archiving, storage & reuse categories of RSpace capabilities illustrating an ecosystem supporting reproducibility and FAIR data. The UCL Research Data Lifecycle, mapping to RSpace electronic lab notebook/ERN, and RSpace infographic of features articulates the many facets of research data management and curation processes which are required in data science as part of support services and the role of libraries.

In order to comprehensively address the data science ecosystem at scale across the institution and all affiliates, partnerships with key stakeholders such as Office of Research, Research Compliance Office, Academic Units, Faculty, IT, Deans, Directors and more are recommended. Institutional-wide research data management policy and ERN solution are highly recommended. Additionally, structuring research through the use of standardize protocols software applications such as protocols.io (<https://www.protocols.io>) with integration of CRediT (Contributor Roles Taxonomy) author statement (<https://www.elsevier.com/authors/policies-and-guidelines/credit-author-statement>) can improve reproducibility, replicability, and FAIR principles in data science.

<sup>6</sup> CMU. (2022). Carnegie Mellon University (CMU) Cloud Lab. <https://cloudlab.mcs.cmu.edu/>.

<sup>7</sup> Smith II, P. L. (2022). Exploring electronic lab notebooks (ELNs) at a R1 institution in the Southeast USA. <https://www.emerald.com/insight/content/doi/10.1108/DLP-02-2022-0013/full/html>.

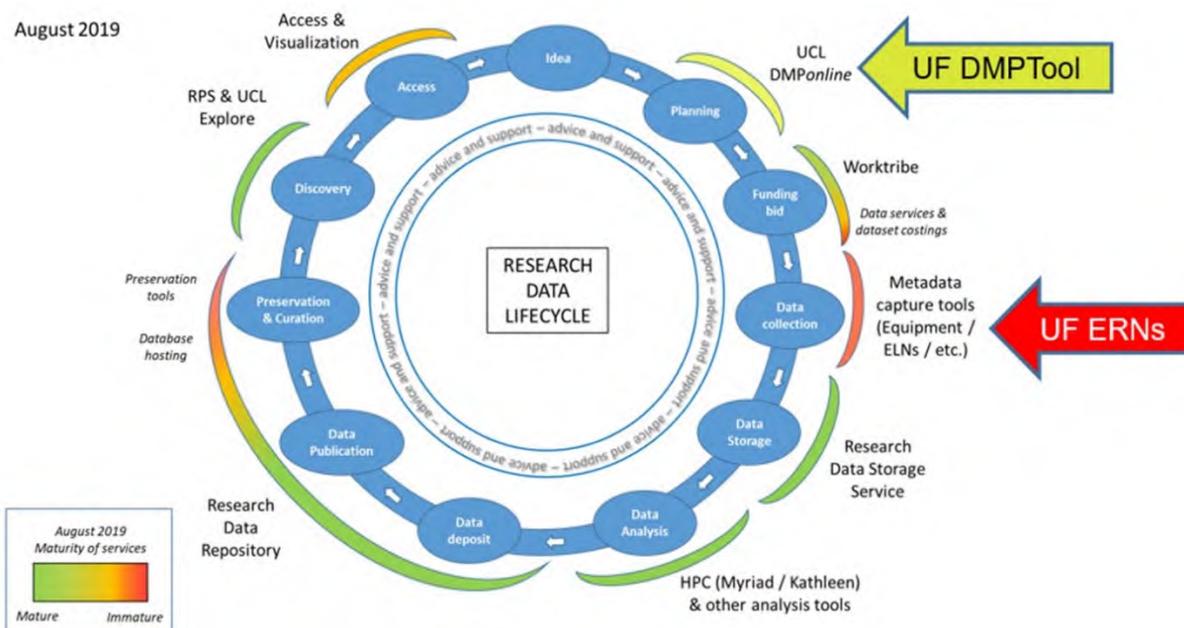


Figure 2. University College London (UCL) Research Data Lifecycle (Wilson, 2022) (used with permission)

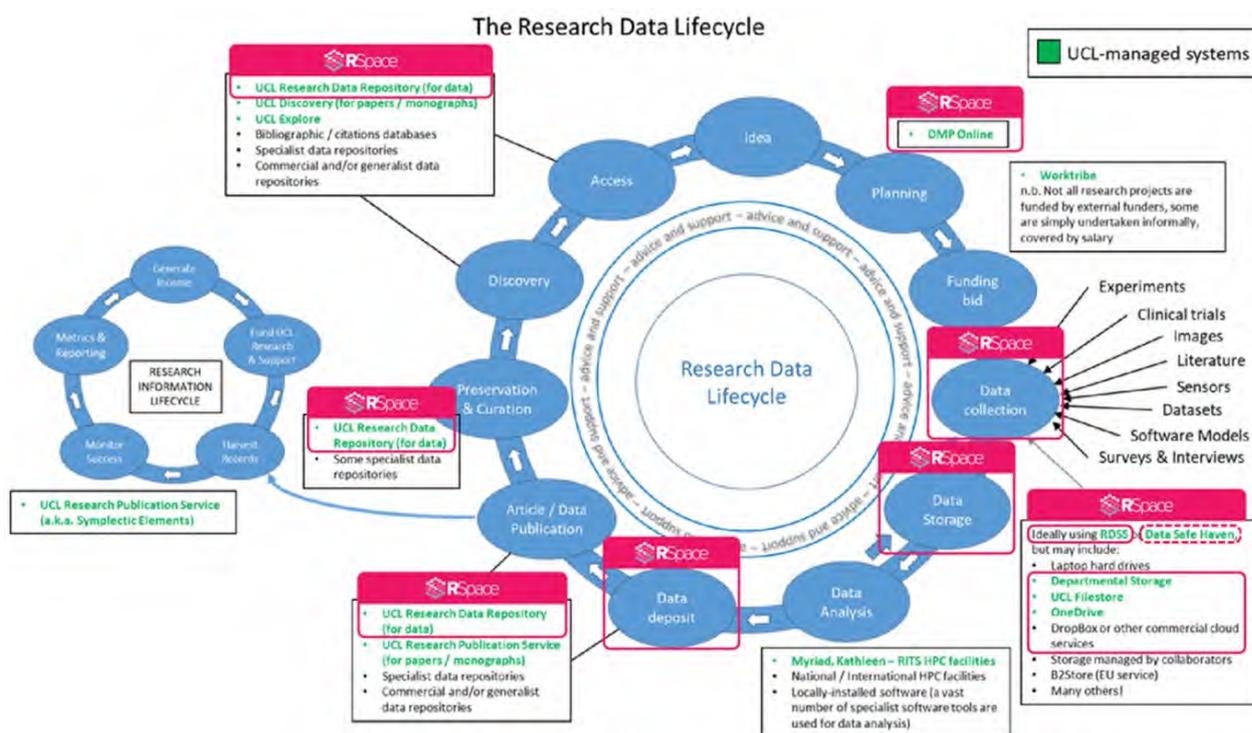


Figure 3. UCL Research Data Lifecycle mapped to RSpace electronic research notebook (ERN) (used with permission - image courtesy of Rory MacNeil, RSpace)

The University College London Research Data Lifecycle (Figure 2) corresponding to maturity level of services aligned with implementation of an institutional ELN solution (Figure 3) illustrates the significance of an ELN capability to address some key research data management (RDM) processes based on selected ELN features and integrations (Figure 3) enabling researchers to meet funders’ evolving RDM demands and data sharing requirements for funded research.

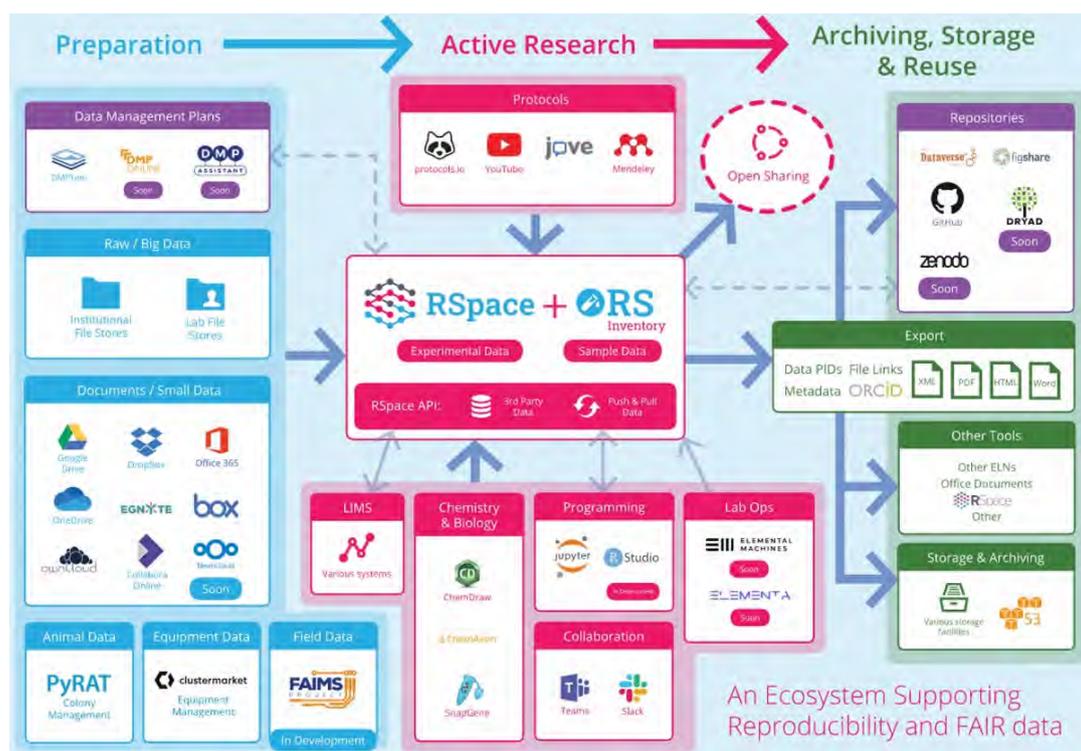


Diagram 1. RSpace Ecosystem Supporting Reproducibility and FAIR data

### CNI Executive Roundtable – Call of Expressions of Interest: Data Science as a Research Support Service and the Role of Libraries

The Libraries led development of the team, expression of interest, and answers to the ten questions for the CNI Executive Roundtable. The majority of the content contribution was from IT Research Computing and Clinical and Translational Science IT Directors. Below are the questions and responses from the authors' institution:

#### 1. How are research support services organized, staffed, supported and funded at your institution?

- i. The university provides subsidy in the form of funding for facilities, electrical power, staff salaries. In addition, some services and staff time are charged in a fee for service model where hourly rates are charged as part of a service center model to recover costs and allow for scaling.
- ii. Some infrastructure is available to all: Dropbox, Google Drive, Microsoft 365 OneDrive
- iii. Large scale computing and data storage infrastructure costs are covered by start-up funds and grants
- iv. Some software licenses are provided for everyone, specialized software must be funded by departments and researchers

#### What is included in the portfolio of research support services?

- i. Access to infrastructure, including infrastructure for regulated research
- ii. Training at various levels of expertise for use of infrastructure, software
- iii. Consulting about writing proposals, data management plans, research workflows and tools

#### How is support services funding related to grant proposals and other sources of funding that directly support researchers?

- i. Basic training is supported by the university
- ii. Grants and start-up packages provide funding for in-depth and advanced consulting and service activities

- iii. University subsidy is spent in devoting time to help develop proposals in collaboration with researchers at no cost to researchers.

**Are costs for institutional support services written into grant budgets?**

- i. The cost for access to research infrastructure and advanced consulting is written into grant budgets
- ii. The cost to execute the implementation and ongoing support of any non-subsidized systems, tools, workflows, data management or statistical support efforts that would directly support a grant is written into a grant budget.

**2. What is the scope of data science research support services?**

- i. All faculty, students and staff and collaborators are supported by the services
- ii. Multidisciplinary teams are supported across colleges and disciplines.
- iii. Multisite and multi university teams are supported.
  - 1) For example, a clinical research study may have 5000 participants at 20 sites at 20 universities around the US with the main data collection and research systems housed at UF and supported by research support teams at UF providing remote support to all 20 site teams.

**3. What is the library's role in research support service provision?**

- i. They provide consulting and training support for students and staff
- ii. There is close collaboration between the libraries and IT for research computing and data management

**Which services does it provide or support? Who are the players?**

- i. Within the library, data science research support services include but are not limited to Academic Research Consulting & Services (ARCS) (<https://arcs.uflib.ufl.edu/>). ARCS research support services include Artificial Intelligence (new/developing), Copyright & Fair Use, Data Management, Data Science & Statistics, Digital Humanities, Publishing & Archiving, Reproducibility, Research Integrity, Research Metrics & Impact, Spatial Information Services, Student HiPerGator Access, Systematic Reviews, and Visualization.
- ii. Students HiPerGator Access ARCS member include: Reproducibility Librarian, Artificial Intelligence Librarian, Bioinformatics Librarian, Geospatial Librarian, and Data Management Librarian (<https://arcs.uflib.ufl.edu/services/student-hipergator-access/>)

**4. How are you staffing this support function?**

**What are the qualifications, backgrounds, and training for support providers?**

- i. In IT research computing the "research facilitators" have scientific and technical backgrounds, often master and PhD degrees

**How much disciplinary expertise (and in which disciplines) is expected along with generalized data science expertise?**

- i. Disciplinary expertise is welcome but not specified, most research facilitators have some strong area, but after a few years on the job, they become generally competent on a broad range of topics

**How large is the staff and how are they organized?**

- i. In IT research computing there are 23 staff organized in teams: DevOps, Training, Services and Support, AI support added last year

**5. How do you define the boundaries, differences, and relationships between research data management and research data science support at your institution?**

Note that at some organizations data science and research data management are treated together and considered as a continuum of activities; in others they are very distinct.

The boundaries, differences, and relationships between research data management and research data science support at the institution originally started as a very distinct separation. However, as research challenges become more interdisciplinary along with research data management requirements for funding agencies such as the NIH SRP P42 data management and analysis core requirement in 2018, research data management is considered as a continuum of activities before, during, and post award throughout the lifecycle of research (<https://research.ufl.edu/research-lifecycle.html>).

**6. Who are the primary users of these research support services? Faculty only? Graduate students? Undergraduates?**

- a) IT research computing support is organized through faculty, either to their research groups or through their teaching of courses
- b) The libraries do offer direct services to postdocs, graduate students, and undergraduates
- c) Clinical and Translational Science Informatics and Technology Service Center offers services to the Health Science community typically.

**7. How much demand is there for these research support services and how are you managing or rationing this demand?**

- i. There is now more demand than capacity to serve

**What are the limits on institutional general purpose research support services in this area? In other words, when do research programs need to shift to dedicated, self-funded data science support staff or collaborators?**

- i. Within the libraries, the limits on institutional general purpose research support services is within the job responsibilities of librarians and staff. However, when the demand for research support services exceed standard job responsibilities, then librarians and staff percentage of time can be included in the budget of grants for institutional support. Faculty librarians are encouraged to seek funding for research support services.
- ii. Key stakeholders (e.g. IT research computing) are trying to develop a business model that will scale to the current demand

**8. What disciplines are making use of, perhaps relying upon, these research support services?**

- i. All 16 colleges make use of these, life sciences (medicine, veterinary medicine, agriculture and food sciences), engineering, liberal arts and sciences are the top users

**Are there other areas (life sciences, for example) where these functions are being handled through funded research collaborations outside the scope of institutional research support services?**

- i. All colleges provide part of the funding through funded research and faculty start-up packages

**9. How are academic programs at the institution organized, how do they relate to research support services (for example, participants in academic programs might supply expertise to research support services capstone projects, clinics, or other mechanisms)?**

- ii. Programs are organized by colleges and institutes
- iii. Programs all use the services in the same way

**10. What training or education in data science is being offered as part of research support services (other than formal academic program courses), and who is offering it?**

- i. The libraries and IT offer special short training opportunities that complement semester-long courses taught by the colleges
- ii. These include: Training videos, training sessions, workshops, multi-session certificate programs, birds-of-a-feather sessions

The libraries will need to continue to develop collaboratively positive and strong working relationships with key stakeholders such as IT research computing, clinical and translation science IT and others to develop research data management and research data science support as a robust sustainable service at scale.

### **Developing socio-technical collaborations with multiple stakeholders**

The development of data science as research support services, particularly the data management and curation components, require social and technical systems collaboration beyond simply addressing software, technology, and tools. A socio-technical systems theory approach address some of the complex, interdependent, and interrelated systems which include technology. Socio-technical systems approach includes “the interrelated nature of t technological and social aspects of the workplace” (Davis et al., 2014, p. 171; Trist and Bamforth, 1951; Trist et al., 1963). A good socio-technical systems approach includes the interrelated nature of goals, people, buildings/infrastructure, technology, culture, and processes/procedures embedded within an external environment of financial/economic circumstances, regulatory frameworks, and stakeholders (Davis et a., 2014, p.173). A socio-technical systems approach enables collaboration.

The Data Management Librarian develops research data management and curation working relationships with multiple stakeholders across diverse and disperse research communities of practice. Three recent international conference presentations all discussed the common theme of developing collaborations with stakeholders across multiple communities of practice on campus. The following three international conference presentations articulates developing collaborations.

- Developing socio-technical collaborations to support researchers at the University of Florida. 42<sup>nd</sup> International Association of University Libraries (IATUL), Miami, USA. Zenodo. <https://doi.org/10.5281/zenodo.6652219>. (2022, June 16).
- Research Support Services and the Role of Libraries in the Lone Cabbage Oyster Reef Restoration Project. 17<sup>th</sup> International Digital Curation Conference (IDCC22), Virtual. Zenodo. <https://doi.org/10.5281/zenodo.6632036>. (2022, June 10).
- Developing data management support partnerships and collaborations at the University of Florida. 47<sup>th</sup> International Association for Social Science Information Service and Technology (IASSIST 2022), Gothenburg, Sweden/Virtual. Zenodo. <https://doi.org/10.5281/zenodo.6618148>. (2022, June 6).

The following non-exhaustive list of stakeholders have collaborated with the libraries on select research support services including data management, sharing, and more.

1. **Office of Research** – collaborated on development and implementation of ORCID API feature integration to enable sharing of award information at UF with ORCID in 2022
2. **Information Technology Research Computing** – collaborated on several projects such as data collection for Association of American Universities and the Association of Public and Land-grant Universities Accelerating Public Access to Research Data meeting in 2020, 2.) Compiled Public Comments on a Draft National Institutes of Health (NIH) Policy for Data Management and Sharing and Supplemental DRAFT Guidance in 2020, 3.)and Office of Science and Technology Policy Request for the Public Comment on DRAFT Desirable Characteristics of Repositories for Managing and Sharing Data Resulting from Federally Funded Research in 2020, and 4.) CNI Executive Roundtable meeting on Data Science as a Research Support Service and the Role of Libraries in 2022).
3. **Informatics Institute** – collaborated on marketing, promotion, and venue for numerous data management training workshops from 2016 to present

4. **Clinical and Translation Science Information Technology** – collaborated on several projects such as those listed under #2 for Information Technology Research Computing
5. Academic Units (select)
  - i. **Department of Anesthesiology** – provided Zenodo data repository training for a lab
  - ii. **Department of Biology** – consulted, recommended, and suggested ELN for faculty
  - iii. **College of Engineering** – consulted on multiple data management plans for faculty
  - iv. **Department of Infectious Diseases and Global Medicine** – provided domain-specific data repository consultation and NIH BMIC Data Sharing Resources<sup>8</sup> training that led to discovery, identification, and selection of NIAID TB Portals<sup>9</sup> for scientist
  - v. **Department of Molecular Genetics & Microbiology** – provided data repository and electronic lab notebook consultations, recommendations, and solutions for faculty
  - vi. **Division of Pediatric Hematology Oncology** – provided data management plan consultation that led to a successful American Brain Tumor Association award
6. Institute of Food and Agricultural Sciences (select)
  - i. **Nature Coast Biological Station** – provided data management plan consultations that led to successful awards leading to development of a new wet-lab in Cedar Keys, FL
  - ii. **Space Life Sciences Lab** – provided numerous reference research articles support
  - iii. **Tropical Research Education Center** – developed an internal successful grant award that funded digitization of research data for development of a digital library resource resulting in development of data archive leveraging Zenodo general data repository ([https://zenodo.org/communities/uf\\_ifas-trec\\_dig\\_lib/](https://zenodo.org/communities/uf_ifas-trec_dig_lib/))
7. **Marston Science Library** – collaborated on the Libraries' 1<sup>st</sup> Data Symposium in 2018 resulting in select presentations archived leveraging Zenodo general data repository (<https://zenodo.org/communities/uf-data-symposium/>)
8. **Center for Environmental & Human Toxicology** – collaborated on developing NIH SRP P42 Data Management and Analysis Core proposals (\$13m; \$12m) in 2018 and 2021
9. **Health Science Center Library** – collaborated on a university-wide ELN survey in 2020 resulting in an ELN Report and Recommendation
10. **UF Scripps Biomedical Research** – provided electronic lab notebook consultations and recommendations in addition to connecting researchers and vendor in 2022
11. **UF Innovate | The Hub** – collaborated with external stakeholders on funding the Libraries' 1<sup>st</sup> data symposium evening reception in 2018

#### Research support drivers (training future researchers)

There is a continuous need to develop data science support training for distinct different groups of learners for undergraduate, graduate/master's, and Ph.D./postdoc. The American Geosciences Institute (AGI) developed a series of career compasses<sup>10</sup> for specialties within geosciences relevant to other disciplines. "Career Compass provides options, tips, suggestions, and strategies for how students can obtain critical skills, experiences, and competencies in order to launch their geoscience career [particularly data sciences]".

Research data management support services (RDMS) are important across all disciplines for good data lifecycle management before, during, and after funded research. Key RDMS

<sup>8</sup> NIH. Trans-NIH Biomedical Informatics Coordinating Committee (BMIC). [https://www.nlm.nih.gov/NIHbmic/nih\\_data\\_sharing\\_repositories.html](https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_repositories.html).

<sup>9</sup> NIH. National Institute of Allergy and Infectious Diseases TB Portals. <https://tbportals.niaid.nih.gov/>.

<sup>10</sup> AGI Career Compasses. <https://www.americangeosciences.org/workforce/compass>.

components include but are not limited to: Data Management Planning, Managing active data, Data selection and handover, Data repositories, Data catalogues, Guidance, training and support, RDM policy and strategy and Business plans and sustainability (Jones et al., 2013).

Future projects related to data support services initiated and led by the libraries at the University of Florida include 1.) Data Services Continuing Professional Education (DSCPE)<sup>11</sup> Capstone Project Partners<sup>12</sup> (Fall 2022) and 2.) Ithaka S+R Coordinating Research Data Support Across Campus<sup>13</sup> (2023-2025). Both projects aim to further develop data management and services support and the role of the libraries.

### **Conclusion and Recommendations**

The libraries must collaborate with partners to offer research computing, Kaggle styled, and Carpentry training courses to help teach and develop foundational coding and data science skills to researchers rather than having librarians and staff develop Carpentry training courses within the libraries. For example, the UF Informatics Institute host the UF Carpentry Club which are organizers of Software and Data Carpentry workshops at the University of Florida (<https://www.uf-carpentries.org/>). Since the organizers are mostly doctoral students in various disciplines with domain expertise, development of courses with actual real world data and use cases is essential in developing relevancy in teaching. However, the turnover of students can produce challenges in teaching sustainability. Thus, collaborative partnerships across campus stakeholders are necessary to share the load of developing new data science workshops, including Library Carpentry (Baker et al., 2016), consistently and effectively across semesters, time, and years despite staff turnover.

“One challenge that remains for the success of data management is the slow adoption of the ELN by research staff, partly because the ELN is still a work in progress, partly because old working habits are slow to change” (Foster et al., 2022; Argento, 2020) [journal article] Foster, E. D., Whipple, E. C., & Rios, G. R. (2022). Implementing an institution-wide electronic lab notebook initiative. *Journal of Medical Library Association*, 110(April), 222-227.

Program evaluation is imperative in the development, growth, and sustainability of research support services. Program evaluation assessment domains include but are not limited to: The need for the program, the design of the program, program implementation and service delivery, program impact or outcomes, and program efficiency Rossi et al., 2004, p.29). Program evaluation assessment of support services include one or more of the five program domains.

---

<sup>11</sup> <https://dscpe.github.io/>

<sup>12</sup> <https://dscpe.github.io/capstone/>

<sup>13</sup> <https://sr.ithaka.org/blog/coordinating-research-data-support-services-across-campus/>

## References

- [journal article] Argento, N. (2020). Institutional ELN/LIMS deployment. *EMBO Reports*, 21(3), 1–8. <https://doi.org/10.15252/embr.201949862>.
- [journal article] Baker, J., Moore, C., Priego, E., Alegre, R., Cope, J., Price, L., ... Wilson, G. (2016). Library carpentry: Software skills training for library professionals. *LIBER Quarterly*, 26(3), 141–162. <https://doi.org/10.18352/lq.10176>
- [website] Carpentries. <https://carpentries.org/>
- [website] CODATA-RDA Schools of Research Data Science. (2022). <https://codata.org/initiatives/data-skills/research-data-science-summer-schools/>.
- [website] Data Science Process Alliance. (2022). <https://www.datascience-pm.com/>.
- [journal article] Davis, M. C., Challenger, R., Jayewardene, D. N. W., & Clegg, C. W. (2014). Advancing socio-technical systems thinking: A call for bravery. *Applied Ergonomics*, 45(2 Part A), 171–180. <https://doi.org/10.1016/j.apergo.2013.02.009>.
- [journal article] Foster, E. D., Whipple, E. C., & Rios, G. R. (2022). Implementing an institution-wide electronic lab notebook initiative. *Journal of Medical Library Association*, 110(April), 222–227.
- [online source] Jones, S., Pryor, G. & Whyte, A. (2013). 'How to Develop Research Data Management Services -a guide for HEIs'. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: <https://www.dcc.ac.uk/guidance/how-guides>.
- [website] Kaggle. <https://www.kaggle.com/learn>.
- [website] Library Carpentry. <https://librarycarpentry.org/>
- [journal article] Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, 1(1), 51–59. <https://doi.org/10.1089/big.2013.1508>.
- [book] Rossi, P., Lipsey, M. W., Freeman, H. E. (2004). *Evaluation: A Systematic Approach*, 7<sup>th</sup> Edition. Sage Publications, Inc. Thousand Oaks, CA.
- [website] Saltz, J. (2022). CRISP-DM is Still the Most Popular Framework for Executing Data Science Projects. <https://www.datascience-pm.com/crisp-dm-still-most-popular/>.
- [book] Schutt, R. K. (2006). *Investigating the Social World*, 5<sup>th</sup> ed. Sage Publications. Pine Forge Press.
- [journal article] Trist, E.L., Bamforth, K.W. (1951). Some social and psychological consequences of the longwall method of coal-getting: an examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. *Human Relations* 4 (1), 3e38. <https://doi.org/10.1177/001872675100400101>.
- [book] Trist, E., Higgin, G., Murray, H., & Pollock, A. (1963). *Organizational Choice (RLE: Organizations): Capabilities of Groups at the Coal Face Under Changing Technologies* (1st ed.). Routledge. <https://doi.org/10.4324/9780203436325>.
- [website] University of Florida Clinical and Translational Science Institute. (2022). Informatics Consulting with CTS-IT. <https://www.ctsi.ufl.edu/research/study-design-and-analysis/informatics-consulting/>.
- [website] University of Florida Libraries. (2022). Academic Research Consulting & Services (ARCS) Student HiPerGator Access. <https://arcs.uflib.ufl.edu/services/student-hipergator-access/>.
- [website] University of Florida Research Computing. (2022). Research Computing. <https://www.rc.ufl.edu/>.
- [website] University of Florida Research Computing. (2022). UFRC Help and Documentation. [https://help.rc.ufl.edu/doc/UFRC\\_Help\\_and\\_Documentation](https://help.rc.ufl.edu/doc/UFRC_Help_and_Documentation).
- [website] University of Florida Research Computing. (2022). UFRC Training. <https://help.rc.ufl.edu/doc/Training>.
- [journal article] van der Voort, H., van Bulderen, S., Cunningham, S., & Janssen, M. (2021). Data science as knowledge creation a framework for synergies between data analysts and domain professionals. *Technological Forecasting and Social Change*, 173(January), 121160. <https://doi.org/10.1016/j.techfore.2021.121160>

## Legal Implications of the CASE Act on Grey Libraries and Grey Literature Authors

Tomas A. Lipinski and Laura C. Schein

School of Information Studies University of Wisconsin—Milwaukee, United States

### Video Presentation

<https://av.tib.eu/media/59871>

### Abstract

*The United States Congress enacted the Copyright Alternative in Small Claims Enforcement (CASE) Act at the end of 2020. The Act created a new Copyright Claims Board (CCB) to hear cases involving smaller claims of infringement as well as noninfringement, in addition to claims of misrepresentation under subsection 512(f) of Title 17, United States Code. The CCB commenced operations in early Summer of 2022. This study investigates the potential impact of the CASE act on library practices with respect to collections of grey literature including the potential for claims to be brought against the library by grey authors. The impact of the CASE Act on the removal and disabling features of Section 512 are examined with the result that libraries may become more engaged in Section 512 processes including the receipt of removal and disabling requests and the halting of restoration mechanisms. The relative infancy of the Copyright Claims Board creates an evolving and uncertain environment in the coming years for the Grey community. Will the CASE Act empower authors of grey content to bring libraries into unwanted CCB proceedings? What protections exist for libraries and their employees in the copyright law and in the CASE Act in specific? Are libraries and patrons exposed to greater legal risk from CASE Act processes and CCB proceedings? What level of resource consumption can be expected in libraries when responding to claims? By examining the mechanisms of the CASE Act, this research can help in understanding where library policies could be strengthened or instigated. Likewise, education resources for both library patrons and grey authors can be formulated.*

*Keywords: Copyright, Copyright Claims Board (CCB), Section 512 Takedown and Disabling, Grey Literature, Library and Archive Policy*

### Introduction

In 2011 Congress requested the United States Copyright Office conduct a study exploring how an administrative process might be instituted to resolve smaller copyright claims; those not exceeding \$30,000.00. The 2013 report “recommended that Congress consider the creation of an alternative forum that will enable copyright owners to pursue small infringement matters and related claims arising under the Copyright Act” (United States Copyright Office, 2013, 3). Since that time, “[t]he record that was developed shows that the concerns that sparked this process more than 15 year years ago remain valid, and that the need for this legislation has increased rather than abated.”<sup>1</sup> The H.R. Report indicated that the mechanism is “intended to be a forum for lower-value copyright disputes in which participation is voluntary for both claimants and respondents... accessible especially for pro se parties and those with little prior formal exposure to copyright laws who cannot otherwise afford to have their claims and defenses heard in federal court.”<sup>2</sup> While litigation against educational entities, libraries and archives is uncommon it does occur. One example is the ongoing litigation involving the course management practices at Georgia State University.<sup>3</sup> A nonprofit library was also the target of litigation in *Hotaling vs. Church of Latter Day Saints*.<sup>4</sup> The case involving infringing reproductions of a literary work in church branch libraries. As the CASE Act is designed for pro se and less well positioned parties, i.e., those with limited financial resources or expertise, is the risk higher that an educational institution or its library or archive might be “sued” under the CASE Act? Yes and no. While it may be easier and less costly to bring a claim using the new process, there may be less incentive to do so as explained below.

**CASE Act: Basics**

The CASE Act is a complex process-oriented piece of legislation, constituting ten new substantive sections of the United States Code.<sup>5</sup> The implementing regulations are extensive.<sup>6</sup>

**Who**

Three Copyright Claims Officers (CCO) constitute the Copyright Claims Board (CCB).<sup>7</sup> Each member must be an attorney with seven years of experience. Two of which have “substantial experience” with copyright infringement claims. Between the two, the substantial experience must consist of representations of both copyright holders and users. The third member must have “substantial familiarity” with copyright law and “experience in the field of alternative dispute resolution” processes.<sup>8</sup> In addition to the CCO there are at least two Copyright Claims Attorneys (CCA) with three years of “substantial experience in copyright law.”<sup>9</sup> Two of the duties of the CCA are to assist the “public with respect to the procedures and requirements of the CCB” and “provide information to potential claimants” regarding obtaining a 17 U.S.C. § 512(h) subpoena relating to the takedown and disabling provisions of section 512 “for the sole purpose of identifying a potential respondent.”<sup>10</sup> These two provisions work in tandem to assist a claimant when an entity is also a service provider under section 512. That entity could be an educational institution, library or archive. Providing assistance to potential claimants may encourage those unfamiliar with court federal court proceedings to pursue remedy through the CCB. The purpose of the subpoena is to identify a student or patron (“subscriber”) that the claimant believes is infringing copyright through posting or linking to their content on the entity’s online platform. When the service provider is a library, issues of patron privacy are also implicated. Many states have statute protecting patron privacy by prohibiting the release of patron records, however most states also have an exception for a court order. A section 512 subpoena issued via 17 U.S.C. § 1503(a)(2)(C) would satisfy the court order exception.

**What**

There are four claims that can be brought before the CCB. First, copyright holders (“claimant”) can bring an infringement claim.<sup>11</sup> Second, users (“claimant” or “respondent”) can seek a declaration of non-infringement.<sup>12</sup> This may be useful to libraries and archives seeking massive digitization projects involving grey collections seeking to have the project deemed lawful under the copyright law. Third, a party can make a claim of misrepresentation under section 512(f).<sup>13</sup> This might occur when a respondent (your patron, student, subscriber, etc.), who is the target of a takedown and disabling notice believes that there is a good faith belief the initial claim contains a misrepresentation relating to the alleged infringing activity. Likewise, a claimant, the sender of a takedown and disabling notice that is in receipt of a counter notification (restoration request) made by your patron, student, subscriber, etc. may believe that request contains a misrepresentation. Finally, a counterclaim seeking damages against a claimant and arising from either 17 U.S.C. § 106 (infringement of an exclusive rights) or 17 U.S.C. § 512(f) (misrepresentation with respect to a claim or restoration/counter notification request) can be filed.<sup>14</sup>

**What Not**

A claim or counterclaim that is not one of the four elucidated claims in section 1504(c)(1)(4) is not permissible.<sup>15</sup> In addition, a claim that has been fully adjudicated or is pending in a court of competent jurisdiction, may not proceed before the CCB unless that court has granted a stay allowing the claimant to proceed to the CCB.<sup>16</sup> Any claim by or against a state or federal governmental entity is also prohibited.<sup>17</sup> This would exclude claims against a state college or university, its library or archive. A claim against a person

or entity residing outside of the U.S is also excluded unless that person or entity initiated the proceeding before the Copyright Claims Board, and remains subject to counterclaims under this CASE Act.<sup>18</sup>

### ***The Impact of the CASE Act on Foreign Parties***

A foreign entity may be subject to CCB jurisdiction if service of process can be established. The individual or entity may be a foreign national but must have a business presence here in the United States. This would be the same for a copyright claim brought in federal district court. However, under the CASE Act one of the four excluded claims include a “claim or counterclaim asserted against a person or entity residing outside of the United States.”<sup>19</sup> In the 231 claims filed through 11/6/2022, all but one misrepresentation case (22-CCB-0020) was dismissed due to non-compliant claims that (a) failed to be amended or (b) addressed a foreign party, or both (22-CCB-0030). An exception is made for those instances where “the person or entity initiated the proceeding before the Copyright Claims Board and is subject to counterclaims under this chapter.” In other words, if a claim is made to the CCB by a person or entity residing outside of the United States and a counterclaim is made by a respondent against the foreign person or entity the claim and counterclaim will be allowed: “A claim or counterclaim asserted against a person or entity residing outside of the United States, except in a case in which the person or entity initiated the proceeding before the Copyright Claims Board and is subject to counterclaims under this chapter.”<sup>20</sup> A library or archive with grey content in its collection located in another country that did not instigate a proceeding under the CASE Act would not be subject to jurisdiction of the CCB under the CASE Act but a library or archive with grey content in its collection that instituted a claim against a U.S. party would be subject to the jurisdiction of the CCB.

### **Methodology**

In addition to reviewing the 2011 U.S. Copyright Report, the statutory text, codified at 17 U.S.C. §§ 1501-1511 was reviewed. To date there have been numerous regulations implementing the statutory text were also reviewed, 37 C.F.R. §§ 220.1 - 234.2. The first claim was filed June 16, 2022.<sup>21</sup>

### **The CASE Act in Operation**

When a claim is filed, the CCA reviews it to determine if it is compliant claim with the statute and applicable regulations. If the claim is determined to be compliant the claimant must file against respondent within 90 days.<sup>22</sup> If claimant fails to then serve the respondent, the CCB dismisses the claim without prejudice. The CCB also notifies the respondent. Either party can request dismissal before the other party responds. If not compliant, the claimant is notified and provided 30 days to perfect the claim. If no filing made within that period CCB dismisses without prejudice. If the claim is refiled but still not compliant; a second notification is given with another 30 days to rectify. If the claim is still not compliant then CCB dismisses without prejudice. If at any time a claim is compliant; the claimant is instructed to proceed. The evidentiary standard used in the review and ultimate determination by the BBC is a preponderance, i.e., the fact asserted is more likely than not to be true. All CCB proceedings qualify as alternative dispute resolution for the purposes of the federal rules of civil procedure.<sup>23</sup>

### **General Opt-out**

Once the claimant notifies a respondent, the respondent has sixty days to decide whether to opt out of the CCB proceedings. All respondents possess this right as the CCB proceedings are voluntary. The CCB can extend the sixty-day period “in the interests of justice.”<sup>24</sup> If a respondent opts out, the CCB dismisses the complaint without prejudice.

However, if the respondent fails to opt out, the respondent cannot later seek remedy in federal court.<sup>25</sup> In other words, failing to exercise the opt out right obligates the respondent to participate in the CCB proceedings. If a case is pending in federal district court and a pending or active proceeding is before CCB involving the parties, the district court “shall issue a stay of proceedings or such other relief as the court determines appropriate.”<sup>26</sup>

### **Specific Library and Archive Opt Out**

While a library or archive with grey content in its collections may take advantage of the voluntary nature of CCB proceedings, a qualifying library or archive may also take advantage of a preemptive opt-out mechanism.<sup>27</sup> The CCB “maintains on its website a public list of libraries and archives that have preemptively opted out.”<sup>28</sup> The Opt Out Registry (the “public list”) for Libraries and Archives is available on the CCB website.<sup>29</sup> As of December 3, 2022, 932 libraries and archives have opted out. To be placed on the “public list” a library or archive must file a certification under penalty of perjury that the library or archive qualifies inclusion.<sup>30</sup> A library or archive that qualifies for the limitations on exclusive rights under 17 U.S.C. § 108(a)(1)-(3) is permitted to be included in the registry if it so desires.<sup>31</sup> The certification should include “any factual statements in support” thereof of its qualification.<sup>32</sup> The certification must be signed by person with “authority.”<sup>33</sup> Here, authority means that the person can take “legally binding action on behalf [of the library or archive] ... in connection with litigation.”<sup>34</sup> This person serves as a point of contact for future correspondence (“including” current phone number, mailing address and email address.<sup>35</sup> For example, if the Register of Copyright determines per section 223.2(a)(3) that a library or archive does not qualify, the point of contact (the person with authority) can “provide evidence supporting its qualification for the exemption within 30 days.”<sup>36</sup> A party may also claim the library or archive does not qualify for the public list (registry). If entity fails to respond or Register determines it does not qualify, then library or archive is either not added to or is removed from the public list. The U.S. Copyright Office may “request additional information...”<sup>37</sup> In reviewing the certification, the CCB accepts facts unless “implausible or [in] conflict with sources ... known to CCB or general public.”<sup>38</sup> If a federal court determines the library or archive does not qualify by meeting the requirements of 17 U.S.C. § 108(a)(1)-(3), the library or archive must notify CCB and provide a copy of the order or opinion within 14 days after determination.<sup>39</sup> A claimant may assert in its claim per 17 U.S.C. § 506(e) and applicable regulations that it believes the library or archive “is improperly included” but the claimant “must include material facts...sufficient to support” that assertion.<sup>40</sup> If CCB determines the “alleged facts sufficient,” the party is instructed to proceed with “service of claim” against the library or archive.<sup>41</sup> Multiple libraries and archives can be included in the same submission.<sup>42</sup>

The Register of Copyright may not charge a fee to preemptively opt out or to require renewal of that designation.<sup>43</sup> A determination for registry qualification “does not constitute a legal conclusion for any other purpose.”<sup>44</sup> Moreover, qualification extends to “employee [of the library or archive] acting within the scope of their employment.”<sup>45</sup> In other words, library and archive employees share in the preemptive opt-out of their library employer. A library or archive “may rescind its preemptive opt-out election” but “may submit no more than one such rescission notification per calendar year.”<sup>46</sup> In spite of the availability of the opt out registry, it may be desirable in a particular case for the library or archive to participate. Participating in the Registry forecloses that choice unless a rescission request is filed but that change in status (removing its name from the public list) can only be made once a year.

### **Damages**

In addition to posing a more simplified claims and counterclaims process than pursuing litigation in federal district court another significant contrast is the range of remedies available. In federal district court the plaintiff can choose between actual damages, such as a loss of profits, including “any profits of the infringer that are attributable to the infringement and are not taken into account in computing the actual damages” or statutory damages.<sup>47</sup> Statutory damages are set by court within a range provided in the statute, \$750.00 up to \$30,000.00 awarded per infringing work infringed, not per infringing copy. In a CCB proceeding an award of actual damages and lost profits in accordance with section 504(b) are allowed subject to CCB award limits. Statutory damages are allowed but are limited to a maximum of \$15,000.00 per work for timely registered works. And \$7,500.00 per work for those works infringed but not timely registered but not to exceed \$15,000.00 in any single proceeding.<sup>48</sup> Regardless of the number claims or counterclaims brought statutory damages are limited to a total of \$30,000.00 “in any single proceeding” exclusive of attorney’s fees and costs, if applicable.<sup>49</sup> The federal district court can also award injunctive relief such as the issuing of cease and desist orders, the seizure and destruction of infringing material, orders disabling access to infringing content. The CCB can order the cessation of infringing conduct, removing, disabling access to or destruction of infringing content.<sup>50</sup> The CCB can also order a party to cease sending notices under section 512, either by the copyright holder to remove or disable content or by a user to restore removed or disabled content, a counter notification.<sup>51</sup> The filing fee in federal district court is \$402.00 that includes a \$52.00 processing fee, whereas it only costs \$100.00 to file with the CCB.<sup>52</sup> The federal district court can award court costs and attorney fees to a prevailing party as well as penalty enhancement for willful violations of \$150,000 per worked in fringed. There is no damage cap on the total damages in a particular proceeding in federal district court. In contrast, unless the claim, counterclaim or defense is brought in bad faith the CCB cannot award costs and attorneys’ fees.<sup>53</sup> If bad faith is present, costs and attorneys’ fees are capped at \$5,000.00 for a party represented by counsel and \$2,500.00 if the party appeared pro se.<sup>54</sup> In “extraordinary circumstances, such as where a party has demonstrated a pattern or practice of bad faith” the CCB may “award costs and attorneys’ fees in excess” of these amounts.<sup>55</sup> Claims of willful infringement are not allowed. The lower monetary recovery allowance may dissuade copyright holders from bringing claims under the CASE Act.

### **Decisions and Appeals**

The CCB will issue a written published decision including the factual and legal basis of determination (including any terms for future conduct).<sup>56</sup> Determination is by majority vote, but the dissenting CCB member can also file written opinion.<sup>57</sup> CCB decisions are published and made available on a “publicly accessible website.”<sup>58</sup> Other than CCB reconsideration or Register review, a final determination precludes re-litigation by any court, tribunal or the CCB.<sup>59</sup> A submission or statement of a party or witness to a CCB proceeding may not be cited before any court, tribunal or the CCB.<sup>60</sup> A determination by the CCB precludes relitigation before any court or tribunal or the CCB and “may be relied upon for such purpose in a future action or proceeding arising from the same specific activity” in order to halt relitigation or re-determination by the CCB, except in the following circumstances.<sup>61</sup> First, a CCB decision does not preclude those “claims or counterclaims not asserted or not finally determined by the CCB” even if between the same parties or same issues.<sup>62</sup> Second, a determination of ownership decided by the CCB may not be relied upon and have no precedential value before a court, tribunal or the CCB.<sup>63</sup> Finally, except for federal district court review any CCB determination may not be cited as “legal precedent” before any court, tribunal or the CCB.<sup>64</sup>

A party may request reconsideration by the CCB within thirty days of a decision.<sup>65</sup> The standard of review is either “a clear error of law or fact or a technical mistake.”<sup>66</sup> This a relatively narrow door of review. If the CCB denies request for reconsideration, the party has 30 days to appeal to the Register of Copyrights.<sup>67</sup> The standard of review used by the register is whether CCB denial was an abuse of discretion by the CCB in denying the reconsideration.<sup>68</sup> Likely, even more narrow a door of review. The Register either denies or remands to CCB with instructions. The amended final determination by the CCB is not subject to further review, except by federal district court. Within ninety days after a CCB final, amended final determination or Register reconsideration, the party may seek order from federal district court vacating, modifying or amending the reconsideration.<sup>69</sup> However, the federal district court intervention is limited to circumstances of “fraud, corruption, misrepresentation or other misconduct” in the determination or where the CCB “exceeded its authority or failed to render a final determination” or where a default determination or determination based on a failure to prosecute was due to “excusable neglect.”<sup>70</sup> Finally, one year after the later of either a CCB final, amended final determination or Register reconsideration, the party can apply to federal district court seeking to reduce the award to judgment. The federal district court can impose “reasonable expenses” including attorneys’ fees incurred by the aggrieved party.<sup>71</sup>

### **Copyright Trolls and Bad Faith**

When the concept of a streamlined alternative to copyright litigation in federal district court was first discussed there was concern that an opportunist copyright holder or entities would abuse the process. Fears of copyright trolls loomed large. Several mechanisms within the CASE Act prevent the over-zealous individual copyright holder, lawyers and law firms or trolling entities from turning CCB filings into a cottage industry. First the U.S. Copyright Office issued a regulation placing limits on the number of filings a party can make in any 12-month period. A claimant is limited to 30 filings, a lawyer is limited to 40 and a law firm is limited to 80 filings in a twelve month period.<sup>72</sup> Moreover, a “proceeding shall count toward the numerical limitation as soon as it is filed, regardless of how the proceeding is resolved, whether it is found to be noncompliant under § 224.1 or unsuitable under § 224.2 of this subchapter, voluntarily dismissed, or fails to become active due to a respondent’s opt-out.”<sup>73</sup> Therefore, a proceeding counts towards the cap as soon as filed regardless of its outcome. No such limitation exists for federal district court filings. If a claim is filed and the respondent opts-out, and the claimant is a troll, the claimant-troll loses the \$100.00 filing fee as the fee is not refundable. Moreover, this filing counts again the 12-month cap on the total number of filings allowable. This is one reason a library or archive may want to forego using the preemptory opt out process and registry of 17 U.S.C. § 1506(aa). Opting out on a case-by-case basis forces the claimant to lose the filing fee plus the filing counts as a strike against the 12-month filing cap. A library or archive that has preemptively opted out does not have the ability impact the claimant in this way. While it may be a nuisance to opt out on a case-by-case basis, nonetheless the library or archive can send a “message” to over-zealous claimants or their representatives. Second, the damage cap limitation applicable to the CCBs prevents the claimant from threatening the exorbitant statutory damages available in 17 U.S.C. § 504(c)(1) or the willful damages in 17 U.S.C. § 504(c)(2) to extort settlements, as would be the case in federal district court. Finally, the CCB may determine in a particular instance that a party acted in bad faith. If bad faith is determined, the CCB can award costs and attorneys’ fees. The CCB will determine conduct is bad faith when the claim, counterclaim or defense is made “for a harassing or other improper purpose, or without a reasonable basis in law or fact.”<sup>74</sup> If this occurs the CCB “shall” unless “inconsistent with the interests of justice, award reasonable costs and attorneys’ fees to the adversely affected party.” The award can be up to \$5,000.00, or if pro se up to \$2,500.00.<sup>75</sup> In extraordinary circumstances, e.g., bad faith “pattern or practice” CCB may “in the interests of justice” make an award in excess of these limits.<sup>76</sup> Likewise, if the claimant is a troll, the troll cannot recover costs or

attorneys' fees even if the troll is the prevailing party. Other pejorative consequences also result. If more than one instance of bad faith is determined in a twelve-month period, the CCB "shall" bar the bad faith actor from filing a claim for next twelve months, plus dismiss all pending cases filed by the bad faith actor. Where the filing is an active proceeding, the respondent must consent in writing to the dismissal, i.e., the respondent library or archive for example may want its right vindicated by the CCB and so would rather proceed with the CCB process.<sup>77</sup> In counting the occurrences of bad faith, the "award of attorneys' fees or costs against an accused participant, pursuant to §232.3, within the prior 12 months shall establish an instance of bad-faith conduct within the requisite time period. The Board may consider other evidence of bad-faith conduct by the accused participant that did not result in an award of attorneys' fees or costs pursuant to § 232.3, including but not limited to, claims that did not proceed because they were reviewed by a Copyright Claims Attorney and found to be noncompliant or where proceedings were initiated but the respondent opted out."<sup>78</sup> The 12-month bar applies to "a legal counsel or authorized representative from participating on a party's behalf."<sup>79</sup> If this occurs, "the Board will consider requests from that party asking the Board to amend the scheduling order or issue a stay of the pending action to allow that party to find other representation."<sup>80</sup>

### ***Impact on Service Provider: Unintended Consequences and Takedown and Disabling Requests***

A library or archive with grey content in its collections may qualify for the service provider protections of 17 U.S.C. § 512. Educational entities also qualify for the service provider protections of Section 512. "The new subsection (j)(1)(B) definition of service provider, for example, includes universities and schools to the extent that they perform the functions identified in new subsection (j)(1)(B)."<sup>81</sup> Section 512 is voluntary, but many education entities, libraries and archives chose to participate in its processes. Section 512 operates as a safe harbor so that an educational institution, library or archive service provider is shielded from consequences (secondary liability) of the infringing conduct of its students or patrons when those students and patrons post or link to alleged infringing content. Such status as a service provider precludes an award of any monetary damages: "damages, costs, attorneys' fees, and any other form of monetary payment" should the copyright holder pursue legal remedy in federal district court.<sup>82</sup> The safe harbor damage remission is a powerful incentive. In return, a service provider is obligated after receiving a conforming notice to "expeditiously [] remove, or disable access to, the material that is claimed to be infringing or to be the subject of infringing activity."<sup>83</sup> To facilitate this process the educational institution, library or archive service provider must name a Designated Agent to receive these notices.<sup>84</sup>

An additional protection in Section 512 offers immunity to the service provider that removes or disables access to content in response to a takedown or disabling request. The immunity is apart from the remission of monetary award for claims of copyright infringement but would shield the service provider from harms where a "subscriber" patron or student claims harm in having their content removed or access to it disabled. A service provider "shall **not be liable** to any person for any claim based on the service provider's good faith disabling of access to, or removal of..." in response to receipt of a takedown or disabling request under Section 512(c)(1)(C).<sup>85</sup> However, the immunity applies only if the service provider "takes reasonable steps promptly to notify the subscriber that it has removed or disabled access to the material."<sup>86</sup> If this occurs the service provider educational entity, library or archive should also send a copy of the notice to the subscriber and inform the subscriber of the counter notification process. A service provider could also consider including a form letter conforming to the requirements for the counter notification under 17 U.S.C. §512(g)(3). This encourages the subscriber student or patron to fill in and return the form to the service provider which in turn this

may trigger a successful restoration request by the subscriber, patron or student. Upon receipt of a conforming counter notification to restore content or cease disabling of access to it, the service provider must “promptly” inform the sender of the takedown or disabling notice that service provider will restore the material or access to it within 10 business days. The service provider then restores material or access to it within 10-14 days, unless the service provider receives notice that sender of the takedown or disabling notice is filing an “action seeking a court order to restrain the subscriber from engaging in infringing activity...”<sup>87</sup> The CASE Act alters this process in two ways. First, a claim filed against a service provider is not considered compliant if the claim does not also affirm that the claimant first notified the service provider per 17 U.S.C. § 512(b)(2)(E), (c)(3) or (d)(3) to remove or disable access to the alleged infringing content, and the service provider failed to remove or disable access to the material expeditiously upon the provision of such notice.<sup>88</sup> This may increase the sending of takedown and disabling notices to grey libraries and archives that have named a Designated Agent to receive such notices. Moreover, in instances where the complaint is not compliant, the CCB will provide the claimant information on serving a proper notice under section 512.<sup>89</sup> This may also increase the filing of takedown and disabling notices against the library or archive service provider.

The second impact of the CASE Act process is that a “claim or counterclaim before the Copyright Claims Board that is brought under subsection (c)(1) or (c)(4) of section 1504, or brought under subsection (c)(6) of section 1504 and that relates to a claim under subsection (c)(1) or (c)(4) of such section, **qualifies as an action seeking an order to restrain a subscriber from engaging in infringing activity** under section 512(g)(2)(C) if (1) notice of the commencement of the Copyright Claims Board proceeding is provided by the claimant to the service provider’s Designated Agent **before the service provider replaces** the material following receipt of a counter notification under section 512(g); and (2) the claim brought alleges infringement of the material identified in the notification of claimed infringement under section 512(c)(1)(C),” i.e., the alleged infringing conduct resulted from the posting of the content by a “subscriber” student or patron.<sup>90</sup> Under the CASE Act then, a second way to halt a restoration of content or access to it would be to file a complaint with the CCB. A far less costly and simpler process than filing an “action seeking a court order to restrain the subscriber from engaging in infringing activity” under 17 U.S.C. § 512(g)(2)(A)-(C) in federal district court. If notice is provided to the service provider’s Designated Agent of the CCB filing before restoration occurs and the notice alleges infringement under subsection 512(c)(1)(C), i.e., posting infringing content, the restoration process is halted.<sup>91</sup> An unintended consequence might be that a claimant files a claim with CCB to halt restoration under subsection 512(g), then having prevented restoration seeks voluntary dismissal, never having intended to seek CCB determination of the matter in the first instance. An unanswered question would be whether this sequence, of filing, halting restoration followed by a dismissal request under 17 U.S.C. § 1506(q)(1) constitutes bad faith conduct under 17 U.S.C. § 1506(y). Recall, the claimant by seek voluntary dismissal anytime before the “respondent files a response.”<sup>92</sup>

### **CASE Act in Operation: 276 Filings to date; December 26, 2022**

#### ***Dismissals***

The extant dismissed cases from the CCB are early indicators of how the tribunal behaves and may continue to behave. By gathering this data, we can see what types of cases are being brought before the CCB, who are the claimants, characteristics of respondents, the reason for dismissal, as well as how the CASE Act impacts foreign parties and interacts globally. This data can inform how grey authors, libraries and archives could function within the small claims landscape. This research was performed in two phases. The first phase of full case analysis was completed November 6th, 2022, and was updated over a month later, December 25th, 2022, to account for new dismissals (Phase 2). By addressing

these dismissed cases in a two-step format, the data was comparable. While both open and closed filings were evaluated, though this study focuses on data from closed dockets.

### **Methodology**

The Case Search Portal on the Copyright Claims Board was utilized to search between open and closed dockets, to navigate individual cases.<sup>93</sup> Each docket number represents an individual case filing, not all cases brought by a claimant, as claimants may re-file or file against multiple respondents. By taking the action of [View] on the Row associated with an individual filing, all documents related to that filing are displayed. In nearly all cases, the initial claim is visible, as well as a clear description of Parties. In the few filings in which the initial claim was not visible, the type of claim was inferred from the Order to Amend Claim, or the Order Dismissing Claim. If a case is dismissed, the Order Dismissing Claim is always public. In most cases, the requests to link parties, as well as the Service Packets and their waivers are restricted to protect the privacy of respondents. The collection of data began by starting with case 22-CCB-0001 and working upward to Case 22-CCB-0276, at first by looking at closed dockets only. Each case was evaluated to learn (a) Type of Claim: Infringement, Misrepresentation, or Claim for a Declaration of Non-Infringement; (b) Date of Dismissal; (c) Foreign or Domestic Claimant or Respondent; (d) Reason for Dismissal; (e) If the case was dismissed due to “Failure to Amend Claim,” the reason for the instruction to amend was evaluated. Some cases had features that made them unique, which we will address in the Breakdown of Claims section; (f) If a dismissed case was filed under misrepresentation under 17 U.S.C 512(f), the Service Provider was notated. (g) Type of Claimant or Respondent; individuals, organizations, brands, or state agencies. These cases were organized into two batches, (1) Nov. 6, 2022 and (2) Dec. 25, 2022; then were merged to generate themes and conclusions about the CCB as of December 25, 2022.

### **Breakdown of Claims**

The following reasons were given for case dismissals in the year 2022. No amended claim was filed in the time allowed. These cases were dismissed because the claim was not amended to accommodate the Copyright Claims Board’s instruction within the time frame. Some examples of these ordered corrections would be issues such as Copyright Registration, Foreign Respondents, Unrelated Respondents, Clarity, Permissible Remedies, Permissible Claims, Legal or Beneficial Ownership, Relief Sought, among many others. Whenever the CCB orders a claim to be amended, they outline exactly what the claimant must fix, using citations from the claimants filing, to make their claim permissible. For some filings there is only one, where others may have many. As discussed, a claimant has two opportunities to file an amended claim successfully, after which the claim will be dismissed if the claim is still found to be non-compliant. The respondent submitted an opt-out notice for this claim within the 60 day opt-out period.<sup>94</sup> These cases were all deemed permissible by the CCB and the respondents were served according to CASE Act procedure. The respondent then chose to opt out. For the filings which had this outcome, there was not an explicit explanation provided by the claimant, and the claim can refile in the future. The claim and amended claim filed in this proceeding, respectively did not comply with the applicable statutory and regulatory requirements for filings before the board. In these instances the CCB then outlines when the claimant amended the claim twice, plus when it was reviewed, and that it was reviewed by a CCB Attorney and passed to the CCB Officer. The Claimant did not file a Proof of Service or Waiver of Service form regarding either respondent within 90 days.<sup>95</sup> These claimants are permitted to refile. The Board did not receive the respondents address by the deadline. Claimants whose cases are dismissed due to this reason may also refile. This makes it possible for claimants to file without the respondents address initially, however the address required within the designated time frame. In this way no claimant is hindered from filing with the CCB. In some instances all parties have agreed in a written

stipulation that the claim be dismissed with prejudice and the CCB dismisses the claim with prejudice. The single settlement among the dismissals as of December 25, 2022 is unique, in that the Claimant and Respondent settled, and the regulations state that the board should dismiss *without* prejudice, however the claimant requested that the case be dismissed with prejudice.

### **Phase 1**

As of November 6, 2022, 51 cases were dismissed out of 231 (22.08%). 20 of these 51 cases were closed because the claim was not amended in the time allowed under U.S.C § 1506 (f)(1)(B). Four of these cases were brought against foreign claimants. In 9 out of 51 cases, the Respondent opted out, ending the CCB case per 17 U.S.C § 1506(i). Six (out of 51) cases were deemed non-compliant, though the copyright holder had amended their claim twice in each filing. These specific claimants are unable to refile with the CCB per U.S.C § 1506(f)(3). In four cases, the CCB did not receive the respondents address by the deadline and were permitted to re-file. Four additional cases were dismissed because the Claimant failed to provide proof of service within 90 days per U.S.C § 1506(v)(1). Those claimants too, were permitted to refile. A singular claimant withdrew their claim with prejudice per 17 U.S.C. 1506(r)(1). 45 of the 51 cases (as of November 6, 2022) were for Infringement under 17 U.S.C. § 1504(c)(1). Two cases were brought before the CCB for both Infringement and Misrepresentation. Five cases were solely for misrepresentation under 17 U.S.C. § 512(f). All but one misrepresentation case (22-CCB-0020) was dismissed due to non-compliant claims that (a) failed to be amended, or (b) addressed a foreign party, or both; as seen in the case 22-CCB-0030. The CCB outlines to the claimant what constitutes a claim of 512 misrepresentation in the "Order to Amend Non-Compliant Claim" document. In the cases of misrepresentation which addressed foreign parties, and that were dismissed because of failure to amend, they had domestic Service Providers, like Google, or YouTube. This trend maintains in Phase two of our study leading into December. In five of these cases, the copyright was not yet registered. Among these misrepresentation cases, there were two instances of persistent copyright holders. This describes claimants who uploaded repeated documents and files after being ordered to cease. In one case they were suspended while the board conducted its compliance review of the claim.

### **Phase 2**

As of December 25, 2022, 78 additional cases had been dismissed since November 6. 46 of these were dismissed because the claim failed to be amended in the time allotted. 10 cases were dismissed as no Proof of Service waiver was filed within the 90 day period. 9 cases were closed because the respondent opted out, and 3 additional cases were closed due to the respondent's address not being provided by the claimant within the time frame. There were some unique cases. For example, one case was closed because the sound recordings in question were made in 1965, the CCB does not have jurisdiction over pre-1972 Sound Recordings according to 17 U.S.C. 1401 and 17 U.S.C. 301(c). Additionally, in case 220-CCB-0055, the claim of infringement was dismissed because the claimant's case had already been seen and dismissed by the Federal District Court. In another case *Richards v. Hasbro, Inc.*,<sup>96</sup> the Claimant's copyright had been refused. In 22-CCB-0092, the Claimant failed to amend their claim to provide evidence that their representation was eligible to represent them before the CCB. These are some examples of unique filings seen by the CCB thus far. 12 cases of the 78 were claims of misrepresentation. Of these cases, two were for both infringement and misrepresentation. One case was for infringement, misrepresentation, and a declaration of noninfringement. The cases which included a claim of misrepresentation were dismissed for a variety of reasons, with the highest amount of reasons for dismissal (7/12) that the claim was not amended. Two of the cases were dismissed because the claim was found to be non-compliant. One case was dismissed because of failure to provide a Proof of Service waiver by the claimant. A case

was dismissed because the Respondent's address was not provided, and an additional, singular misrepresentation case was dismissed because the respondent opted out. This case, 22-CCB-0256, is the most recent case to date.

### **Trends**

When Phase 1 and 2 are combined, several trends are apparent. When the CCB began in June of 2022, there was a higher rate of Respondents opting out; in fact, out of the first fifteen cases the CCB ruled in June and July, six of these were opt outs. This is over 25% of cases being dismissed for this reason. In contrast, of the most recent fifteen cases dismissed as of December 25, 2022, only one case was due to the respondent opting out of proceedings. The rest of the cases were because the claim was not amended. This reason (claim not amended) is still the leading reason for case dismissals, however the percentage goes up from 39% of dismissals in Phase 1, to 58% in Phase 2. Of all the cases seen by the CCB, most misrepresentation cases seen by the CCB had foreign respondents. These are situations in which either party, or both, are outside of the United States; in which the infringement occurred via a domestic service provider. The misrepresentation was claimed in the response to the service provider's answer to the take down notice. In isolated cases, the work was registered in the United States, even though both were foreign parties. Some countries which are cited in claims are Ukraine, Kazakhstan, Russia, China, Germany, Ireland, and the United Kingdom.

The cases seen by the CCB vary in scale. Most dismissals are individuals filing against other individuals or smaller companies, in many cases, web-based platforms or music recording/distribution entities. There were several cases featuring larger corporations such as Paramount Pictures, Wells Fargo, and Hasbro. Some platforms were also named in filings such as YouTube or Amazon. There were increased filings regarding social media in the Phase two portion of this study. Assessing these dismissals is relevant because it allows us to look comparably at open dockets. After becoming familiar with the closed dockets and their contents, many parties are recognizable. It appears that the feature of the CCB which permits claimants to refile is being utilized. A notable number of individuals who did not amend their claim in an earlier filing or did not provide proof of service within the 90 day opt out period refiled, and these cases are again active proceedings before the CCB. There have been no monetary awards or damages awarded by the CCB as of December 26, 2022. This indicates that to date all dismissals are caused from legal processes versus the verdict of the case itself. As of December 26, 2022, there are no cases involving libraries or archives. While the name of one claimant has "Archive" in the title, it is a photography business platform v. an institutional entity.

### **Dismissals: Implications for Libraries and Grey Authors**

Implications for grey authors will become clearer in the coming months, as the CCB platform becomes more well known. Because the cases which have been dismissed so far have been due to legal thresholds within the CCB process, it is important that grey authors and grey libraries are aware of procedure and boundaries of the CCB. An important notation for grey authors, especially for foreign parties, is to ensure the infringed work has a copyright in the U.S., or is in the process of registration for one. Additionally, if instructed to amend an existing claim, ensure this is done within the designated time, and that the recommendations by the CCB are followed. They instruct the claimant how to properly amend their claim in their "Order to Amend Non-Compliant Claim," and uses direct examples and citations from the filing itself to guide the claimant. Due to the volume of cases being seen by the CCB, it is likely that libraries and archives will be brought into proceedings in the future. The process by which the CCB is streamlining their dismissals indicates that these proceedings will likely be worked out using the respondent opt out option or through the progression of the CCB processes before ever reaching a final ruling. The coming year will indicate how many cases make it to the final stages in the CCB Claims Board processes.

### **Recommendations**

As of December 26, 2022, one third of the preemptive opt out list is comprised of public libraries. The majority of libraries and archives that select this this opt out option are academic in nature. This indicates that public libraries are less likely to be using the preemptive opt out list, making them vulnerable to unwanted CCB proceedings by grey authors. The preemptive opt out option can help prevent libraries from being pulled into initial proceedings and prevent the entity from being involved in repeated filings before the CCB. The CASE ACT as implemented by the CCB makes it simple to file an initial claim; in this way it does expose libraries and patrons to increased legal risk. However, the maintenance of these filings is such that the likelihood of sustained involvement with the CCB is lowered. If libraries and archives are brought into proceedings, and they do not Opt-out, there would be a risk of increased resource consumption; libraries might need to seek legal counsel or make changes to current procedures, or perhaps seek a Declaration of Non-Infringement. With the presence of streaming platforms increasing in library environments (both public and academic), it is increasingly important that libraries and librarians are aware of the CASE Act and its accessibility, as well as threats to patron privacy through increased use of the Section 512(h) subpoena to identify the subscriber-patron. Having a designated employee responsible for handling interactions with copyright is one way that libraries can prepare for these challenges. Increasing efforts to maximize patron privacy through tech based solutions, beefing up cybersecurity efforts and taking proactive measures to protect patron data is key. Some libraries may find it in their best interests to hold programming events to inform their community about how they can protect themselves as well. This could be especially useful in environments which support grey authors. Many of the cases which cited a section 512 misrepresentation, did so in response to the claimants receiving an section 512(c) takedown or disabling notice. Libraries, service providers and rightsholders can attempt to work together when handling takedown notices and restoration services. This can help in furthering the efficacy of the CASE Act by supporting functionality.

Grey authors are at increased risk of being pulled into CCB proceedings. This is partly due to the ability to easily file a claim. The chance that grey authors will be able to preemptively opt out of proceedings is low, and opting-out of proceedings after the claim is initiated puts them at risk for being involved in the Federal District Court proceedings later, which may not be auspicious for independent authors and smaller grey organizations. It is important that grey authors have a comprehensive understanding of copyright procedures, as well as the structure of the CASE Act.

### **Educational Resources for Libraries and For Grey Authors.**

A direct resource for Libraries and Grey Authors needing information about the CCB, can go directly to their Handbook.<sup>97</sup> The Copyright.gov website is also invaluable for both Libraries and Grey Authors. Not only does it outline the laws in place, but it also details changes and updates relevant to the Copyright Claims Board. The American Library Association (ALA) website is also home to valuable resources for library administration, librarians, and patrons.<sup>98</sup> ALA provides comprehensive resource guides to valuable physical and digital resources. The Association of College and Research Libraries,<sup>99</sup> and the Association of Research Libraries<sup>100</sup> are home to valuable libguides and resources for librarians.

### **Conclusion: Benefits**

While the CASE Act allows any respondent to opt-out and for a library or archive to preemptively opt-out there are depending on circumstances benefits to participating in CCB processes. First, a library or archive may seek a declaration of noninfringement that can include a determination of fair use or other lawful uses such as Section 108 (library

and archives) or Section 110 (educational entity). For the nominal \$100.00 filing fee, a use can be deemed a fair use or otherwise non-infringing by the CCB. As the CCB must consider other provisions of the copyright law (“this title”), i.e., copyright statutes and regulations and existing judicial precedent. The CCB must follow the law in federal jurisdiction where copyright holder could have filed in federal district court. If more than one jurisdiction possible, CCB determines which venue provides the connection to the parties and alleging infringing conduct. It is unclear whether this precludes CCB from consulting persuasive precedent; persuasive precedent that may favor a use by the library or archive? The CCB proceedings are designed with the untrained claimant in mind as lawyers, in-person appearances are not necessary, nor is formal motion practice allowed. The CASE Act provides a less formal and costly procedure to secure a determination of noninfringement as parties can proceed pro se or be represented by law student pro bono. On the other hand, attorneys’ fees and costs are awarded only in the presence of bad faith. But this may prove useful when the use is likely to be found infringing by the library or archive. If the library or archive is the prevailing party and the other party fails “to pay damages or otherwise comply with the relief, the federal district court “shall impose...reasonable expenses required to secure such order including attorneys’ fees that were incurred by aggrieved party” in an application to reduce an award to judgment.<sup>101</sup> Moreover, statutory damages are capped at \$15,000.00 per work for registered works and for untimely registered works: \$7,500.00 per work but not more than \$15,000.00 in a proceeding. If the work is timely registered the damages in a single proceeding are limited to \$30,000.00, exclusive of attorney’s fees and costs. Finally, misrepresentation claims under 17 U.S.C. § 512(f) can be brought before the CCB to combat overly aggressive copyright holders and copyright trolls.

### **Future Report to Congress**

There may be changes to the CASE Act as within three years after first decision of the CCB, Congress requested that the Register of Copyrights (“shall”) submit a report addressing the following: the use and efficacy of the CCB, including the number of proceedings the CCB could reasonably administer; whether any adjustments to CCB authority (expansion or limitations of its scope); increasing the eligible claims that could be filed, e.g., 17 U.S.C. § 1202 (removal of copyright management information); adjustments to the “works and applicable damages limitations”; whether greater allowance (or limitations) of attorneys’ fees and costs should be made; adding mechanisms to assist copyright holders to identify and locate online infringers, hopefully including the impact of increased use of 17 U.S.C. § 512(h) (subpoena powers); whether CCB duties should be expanded to offer mediation or alternative dispute resolution services; and any other matter pertinent to the CCB.<sup>102</sup> In addition, the CCB within 3 years of its first determination may “conduct a rulemaking to adjust the limits on monetary recovery or attorneys’ fees and costs.”<sup>103</sup> This rulemaking is likely to include opportunity for public comments. As this element is one of the aspects of the three-year report to Congress it is unclear whether the Register will seek public comment on the other aspects of the required report.

### **References**

- Cambridge University Press v. Albert*, 906 F.3d 1290 (11th Cir. 2018).
- Code of Federal Regulations: 37 C.F.R. §§ 220.1 – 234.2.
- Copyright Alternative in Small Claims Enforcement Act of 2019, Pub. L. 116–260, 134 Stat. 2176 (2020).
- Hotelling vs. Church of Latter Day Saints*, 118 F.3d 199 (4th Cir. 1999)
- H.R. Rep. No 105-551, pt. 2 (1998).
- H.R. Rep. 116-252 (2019).
- S. Rep. 105-190 (1998).
- S. Rep. 116-105 (2019).
- United States Code: 17 U.S.C. §§ 1501-1510.
- United States Copyright Office, Copyright Small Claims, A Report of the Register of Copyrights, 2013.
- United States Code: 17 U.S.C. § 1401.

## Notes

- 
- <sup>1</sup> H.R. Rep. No. 116-252, at 19 (2019).
- <sup>2</sup> H.R. Rep. No. 116-252, at 17 (2019).
- <sup>3</sup> *Cambridge University Press v. Albert*, 906 F.3d 1290 (11th Cir. 2018).
- <sup>4</sup> *Hotaling vs. Church of Latter Day Saints*, 118 F.3d 199 (4th Cir. 1999).
- <sup>5</sup> 17 U.S.C. §§ 1501-1510.
- <sup>6</sup> 37 C.F.R. §§ 220.1 – 234.2.
- <sup>7</sup> 17 U.S.C. § 1502(b)(1).
- <sup>8</sup> 17 U.S.C. § 1502(b)(3)(A)(i)-(iii).
- <sup>9</sup> 17 U.S.C. § 1502(b)(3)(B).
- <sup>10</sup> 17 U.S.C. § 1503(a)(2)(B) and (C).
- <sup>11</sup> 17 U.S.C. § 1504(c)(1).
- <sup>12</sup> 17 U.S.C. § 1504(c)(2).
- <sup>13</sup> 17 U.S.C. § 1504(c)(3).
- <sup>14</sup> 17 U.S.C. § 1504(c)(4).
- <sup>15</sup> 17 U.S.C. § 1504(d)(1).
- <sup>16</sup> 17 U.S.C. § 1504(d)(2).
- <sup>17</sup> 17 U.S.C. § 1504(d)(3).
- <sup>18</sup> 17 U.S.C. § 1504(d)(4).
- <sup>19</sup> 17 U.S.C. § 1504(d)(4).
- <sup>20</sup> 17 U.S.C. § 1504(d)(4).
- <sup>21</sup> Docket 22-CCB-0004.
- <sup>22</sup> 17 U.S.C. § 1504(g).
- <sup>23</sup> 17 U.S.C. § 1509(b).
- <sup>24</sup> 17 U.S.C. § 1506(i).
- <sup>25</sup> 17 U.S.C. § 1504(g)(1).
- <sup>26</sup> 17 U.S.C. § 1509(a).
- <sup>27</sup> 17 U.S.C. § 1506(aa).
- <sup>28</sup> 37 C.F.R. § 223.2(b)(1).
- <sup>29</sup> See, <https://ccb.gov/libraries-archives-opt-out/>
- <sup>30</sup> 37 C.F.R. § 223.2(a)(1).
- <sup>31</sup> 17 U.S.C. § 1506(aa)(4).
- <sup>32</sup> 37 C.F.R. § 223.2(b)(2).
- <sup>33</sup> 37 C.F.R. § 223.2(a)(2).
- <sup>34</sup> 37 C.F.R. § 223.2(c).
- <sup>35</sup> 37 C.F.R. § 223.2(a)(2).
- <sup>36</sup> 37 C.F.R. § 223.2(b)(1).
- <sup>37</sup> 37 C.F.R. § 223.2(b)(1).
- <sup>38</sup> 37 C.F.R. § 223.2(a)(3).
- <sup>39</sup> 37 C.F.R. § 223.2(a)(4).
- <sup>40</sup> 37 C.F.R. § 223.2(b)(2).
- <sup>41</sup> 37 C.F.R. § 223.2(b)(2).
- <sup>42</sup> 37 C.F.R. § 223.2(b)(2).
- <sup>43</sup> 17 U.S.C. § 1506(aa)(3)(A) and (B).
- <sup>44</sup> 37 C.F.R. § 223.2(b)(1).
- <sup>45</sup> 37 C.F.R. § 223.2(a)(5).
- <sup>46</sup> 37 C.F.R. § 223.2(a)(7).
- <sup>47</sup> 17 U.S.C. § 504(b) and (C).
- <sup>48</sup> 17 U.S.C. § 1504(e)(1)(A).
- <sup>49</sup> 17 U.S.C. § 1504(e)(1)(D).
- <sup>50</sup> 17 U.S.C. § 1504(e)(2)(A)(i).
- <sup>51</sup> 17 U.S.C. § 1504(e)(2)(A)(ii).
- <sup>52</sup> 37 C.F.R. § 201.3(g)(1)(i) and (ii).
- <sup>53</sup> 17 U.S.C. § 1504(e)(3).

- 
- <sup>54</sup> 17 U.S.C. § 1506(y)(2).
- <sup>55</sup> 17 U.S.C. § 1504(y)(2)(B).
- <sup>56</sup> 17 U.S.C. § 1506(t)(1)(B) and (C).
- <sup>57</sup> 17 U.S.C. § 1506(t)(1)(A) and (2).
- <sup>58</sup> 17 U.S.C. § 1506(t)(3).
- <sup>59</sup> 17 U.S.C. § 1507(a).
- <sup>60</sup> 17 U.S.C. § 1507(c).
- <sup>61</sup> 17 U.S.C. § 1507(a).
- <sup>62</sup> 17 U.S.C. § 1507(a)(1).
- <sup>63</sup> 17 U.S.C. § 1507(a)(2).
- <sup>64</sup> 17 U.S.C. § 1507(a)(3).
- <sup>65</sup> 17 U.S.C. § 1506(w).
- <sup>66</sup> 17 U.S.C. § 1506(w).
- <sup>67</sup> 17 U.S.C. § 1506(x).
- <sup>68</sup> 17 U.S.C. § 1506(x).
- <sup>69</sup> 17 U.S.C. § 1508(c).
- <sup>70</sup> 17 U.S.C. § 1508(c)(1)(A)-(C).
- <sup>71</sup> 17 U.S.C. § 1508(a).
- <sup>72</sup> 37 C.F.R. § 233.2(a)(1)-(3).
- <sup>73</sup> 37 C.F.R. § 233.2(a).
- <sup>74</sup> 17 U.S.C. § 1506(y)(2).
- <sup>75</sup> 17 U.S.C. § 1506(y)(2)(A).
- <sup>76</sup> 17 U.S.C. § 1506(y)(2)(B).
- <sup>77</sup> 17 U.S.C. § 1506(y)(3).
- <sup>78</sup> 37 C.F.R. § 232.4(c).
- <sup>79</sup> 37 C.F.R. § 232.4(d)(1).
- <sup>80</sup> 37 C.F.R. § 232.4(d)(3).
- <sup>81</sup> H.R. Rep. 105-551, pt. 2, at 64 (1998); Sen. Rep. 105-190, at 54-55 (1998).
- <sup>82</sup> 17 U.S.C.A. § 512(k)(2).
- <sup>83</sup> 17 U.S.C.A. § 512(c)(1)(C).
- <sup>84</sup> 17 U.S.C.A. § 512(c)(2).
- <sup>85</sup> 17 U.S.C. § 512(g).
- <sup>86</sup> 17 U.S.C. § 512(g)(2)(A).
- <sup>87</sup> 17 U.S.C. § 512(g)(2)(A)-(C).
- <sup>88</sup> 17 U.S.C. § 1506(f)(1)(C)(i).
- <sup>89</sup> 17 U.S.C. § 1506(f)(1)(C)(ii).
- <sup>90</sup> 17 U.S.C. § 1507(d).
- <sup>91</sup> 17 U.S.C. § 1507(d).
- <sup>92</sup> 17 U.S.C. § 1506(q)(1).
- <sup>93</sup> See, <https://dockets.ccb.gov/search/cases>.
- <sup>94</sup> 17 U.S.C § 1506(i).
- <sup>95</sup> 17 U.S.C. § 1506(v)(1).
- <sup>96</sup> Docket 22-CCB-0084.
- <sup>97</sup> See, [www.CCB.gov/handbook](http://www.CCB.gov/handbook)
- <sup>98</sup> See, <https://libguides.ala.org/copyright>
- <sup>99</sup> See, <https://acrl.libguides.com/scholcomm/toolkit/copyright>
- <sup>100</sup> See, <https://www.arl.org/know-your-copyrights/>
- <sup>101</sup> 17 U.S.C. § 1508(a).
- <sup>102</sup> Pub. L. 116–260 (CASE Act), § 212(e)(1)-(6) (2020).
- <sup>103</sup> 17 U.S.C. § 1510(a)(2)(A).



# WorldWideScience.org

*An International Partnership Supporting Open Science*



- ▶ Simultaneously explore over 100 national and international scientific databases and portals from 75 countries
- ▶ Search information in textual, multimedia, software, and scientific data formats
- ▶ Eliminate language barriers through multilingual translations across ten languages

## Search and Access Scientific Research Data

**Data tab identifies results from scientific research data collections**

**Data can be viewed or downloaded**



## Modern Science Demands Reproducibility and Open Access to Publications, Datasets, and Software



Operating Agent:

**OSTI.GOV**

## Using the Overton policy to academic citation network: How does the policy grey literature and scholarly record connect?

Euan Adie, Terrence Bucknell, Jennifer Glover, and Ángel Luis Jaso-Tamame  
Open Policy Ltd., United Kingdom

### Video Presentation

<https://av.tib.eu/media/59874>

### Abstract

*As part of the broader impact agenda at universities in knowledge economies (and the UK, Australia and Netherlands in particular) there is growing interest in linking academic research outputs to the policy documents – government guidelines, policy briefs, white papers, impact assessments and so on – that cite them. Using Overton, a novel grey literature database of 6.9M+ full text policy documents from governments and think tanks around the world, we find while in scholarly literature the sciences tend to be highly cited, and the social sciences less so, the opposite is true in the policy literature. We discuss how accessing and organizing the policy grey literature could help universities, funders and governments better understand the impact their outputs are having on the wider world.*

### Introduction

Evidence-based policy is an approach to policymaking that emphasizes a more systematic use of data, research, and scientific evidence in the policymaking process, in an effort to ensure that policies are based on the best available evidence and guided by the principles of scientific rigor and objectivity (Baron, 2018) (European Parliamentary Research Service, 2021).

Its growth since the end of the late 20th century is partially behind what is sometimes called the “impact agenda” in modern research institutions, especially those in the UK, Australia and the Netherlands (Smith, 2020). The impact agenda aims to recognize and reward the impact beyond academia, to encourage research that demonstrably contributes to society and the economy (UKRI, 2022).

One way that research can achieve this is to be used by policymakers who help to enact change through legislation and guidelines. For example, climate change research used by groups like the IPCC contribute to findings that go on to have a direct effect on greener policies and climate change legislation worldwide (Intergovernmental Panel of Climate Change, 2013).

Finding and recognizing what research has been used in policy is difficult (Newson, 2018). Previous approaches have looked for academic style references in policy documents harvested from a relatively small number of websites (Altmetric, 2021) (Wellcome Trust, 2020), but policy documents are frequently transient (e.g. government departments change name, agencies redesign their website), use unorthodox citation styles and are spread across a large number of websites.

### Overton

To overcome some of these challenges we have created a large, continuously updated database of policy related grey literature, named Overton. Available at [www.overton.io](http://www.overton.io), as on January 2023 this database contained 6.9M+ documents collected from 180 different countries. Using the database and associated web interface requires a commercial subscription to be used for researcher assessment purposes, but is free for use in academic research.

Users can search the metadata and full text of policy documents or filter them by year, topic, author, source and more.

Overton differs from previous policy related systems in how it defines and collects documents, finds references and connects other publicly available datasets of scholarly and institutional metadata.

### **Defining what a “policy document” is**

The simplest definition of a policy document might be “documents written by a policymaker”, encompassing a broad swathe of governmental outputs.

However, Overton’s primary use case involves finding all of the links between research and policy, and this definition doesn’t allow us to do that.

Typically these links are not direct: that is, policymakers do not often find and cite research directly. More often research reaches policy via a third party knowledge broker who sits between the worlds of academia and policy (MacKillop, 2019).

Examples of these brokers include think tanks, intergovernmental organizations like the World Health Organization, policy orientated non-governmental organizations and charities like Greenpeace or the ACLU and technical agencies inside government.

The job these brokers do is to collect and assess evidence on a topic and then synthesize it for an often time poor, generalist policymaker audience.

As such in order to trace the pathway from research to policy for any given scholarly book or article Overton must know about outputs from any knowledge brokers as well as the government sources responsible for driving the actual legislation.

Therefore the definition we use is not “documents written by policymakers” but “documents written primarily for or by policymakers”, and the sources we collect from include the categories of organization listed above.

### **Collecting documents at scale**

Overton collects documents at a larger scale than similar systems, fetching from 1,600+ distinct websites that host data from around 31,000 organizations – some websites being aggregators for multiple agencies, groups or departments (Overton, 2023).

A combination of web scraping, data feeds and APIs are used. Overton only collects public documents, that is documents that aren’t behind a paywall or that otherwise require a login. Only documents hosted online are collected: the database has no facility to import paper documents or documents that aren’t accessible through the public world wide web.

Though Overton only hosts document metadata and doesn’t serve full text – users of the system are referred back to the original source if they want to download PDFs or view the full text of pages – the collection system picks up license information where available for use by researchers who may later want to analyse or download the data in bulk.

Unfortunately, it is difficult to assess how complete Overton’s coverage is as there is no existing gold standard to compare against. In some countries with a centralized government document repository – like the UK and Canada – we can be satisfied that we are collecting the majority of available documents by comparing what is available in the system with what is available on those repositories, but in others there is no way to know what is being missed.

### **Finding references**

Overton parses each document it collects to look for references – either explicit links or in text - to scholarly books & articles (defined as anything that registered with CrossRef or DataCite, the DOI registration agencies used by the vast majority of scholarly publishers)

as well as to other policy documents (defined as other documents in the Overton database).

By necessity Overton's definition of what a reference might look like has to be broad. A scholarly reference may contain page, volume and issue numbers, a journal name or abbreviation and follow a known format specified by the document's publisher. They are typically grouped in a bibliography at the end of a chapter or article.

References in policy documents don't follow these patterns. Citations of other policy documents may include as little as just an organization name and title.

Furthermore, context is very important: for example, a policy document in the US may reference a document from the "Department of Transportation" that may be referring to the local department (at either state or federal level) or in extremis the department of transportation in some other country entirely. Referenced titles may be in their original language, or have been translated into the language of the citing policy document.

Though some policy documents do have a bibliography, more often references can be found in footnotes, table captions and in the body of the actual document.

As such Overton looks for a wider range of reference types in more sections of the document, and then includes context specific heuristics to deal with ambiguous organization names and titles.

### ***Analysis of the research / policy citation network***

The current Overton database contains ~ 6.9M documents, from which we have extracted 16.3M references to 4.4M distinct scholarly books and articles, and 6M further references to other policy documents collected by Overton. A detailed, bibliometric centred analysis can be found in Szomszor et al. (Szomszor, 2022).

The ten scholarly articles most cited by policy documents in Overton are presented in Table 1.

<b>Title of article</b>	<b>Journal</b>	<b>Date</b>	<b>Citations</b>
The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity	Econometrica	Nov 2003	2358
Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations	The Review of Economic Studies	Apr 1991	2277
Initial conditions and moment restrictions in dynamic panel data models	Journal of Econometrics	Nov 1998	1633
The Skill Content of Recent Technological Change: An Empirical Exploration	The Quarterly Journal of Economics	Nov 2003	1577
Sample Selection Bias as a Specification Error	Econometrica	Jan 1979	1501
Measuring inconsistency in meta-analyses	BMJ	Sep 2003	1459
The Colonial Origins of Comparative Development: An Empirical Investigation	American Economic Review	Dec 2001	1431
<b>Endogenous Technological Change</b>	Journal of Political Economy	Oct 1990	1395
Staggered prices in a utility-maximizing framework	Journal of Monetary Economics	Sep 1983	1383
On the mechanics of economic development	Journal of Monetary Economics	Jul 1988	1328

*Table 1. The ten scholarly articles cited most often in the Overton policy document database.*

The ten articles in Table 1 are for the most part related to economics: the most cited paper, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity" is heavily cited in scholarly literature too (Google Scholar reports 17,865 citations) and establishes a link between trade liberalization and productivity.

Several are methodological papers – including the single paper from medicine ("Measuring inconsistency in meta-analyses" from the British Medical Journal).

To explore which fields were represented in the full set of papers cited by policy we mapped the journal in which each one was published to the All Science Journal Classification Codes (ASJC) as used by Elsevier’s Scopus platform (Scopus, 2023).

The top 15 subject areas are presented in Table 2, and shows that while economics papers are heavily cited, other subjects are equally well represented, with public health having the highest number of unique articles cited at least once by policy in the database.

One surprising difference to what is typically seen in the scholarly literature – where typically the sciences and especially the “hard” sciences are heavily cited, and social sciences much less so (Patience, 2017) - is that sociology, psychiatry and education are all represented in Table 2.

In fact to some extent the converse is also true: public health and ecology are represented in the top policy cited subject areas, but notable by their absence are the hard sciences: physics, chemistry and electrical engineering don’t appear. Intuitively this makes sense: policy documents aren’t usually about high energy physics, and so don’t cite high energy physics papers.

Subject area	Unique articles cited in policy
<b>Public Health, Environmental and Occupational Health</b>	201,639
Economics and Econometrics	197,704
Sociology and Political Science	155,841
Ecology, Evolution, Behavior and Systematics	143,753
Psychiatry and Mental Health	130,566
Medicine (all)	119,639
Infectious Diseases	115,873
Ecology	105,643
Aquatic Science	101,443
Geography, Planning and Development	97,551
Education	95,028
Oncology	94,243
Pediatrics, Perinatology and Child Health	89,964
Neurology (clinical)	81,352
Management, Monitoring, Policy and Law	81,105

Table 2. Top 15 subject areas by the number of unique articles cited in policy.

## Conclusions and future work

Overton allows users a quick and easy way to search the policy related grey literature from governments and knowledge brokers. By automatically identifying and extracting links between research and policy it also opens up new avenues of investigation into how evidence is sourced and used by policymakers and their intermediaries.

Many challenges remain. Ideally it would be possible to determine how much of the policy literature Overton actually contains, perhaps by looking for references in the collected set that haven’t been matched to anything else already in the database, or by performing manual spot checks in different countries and topic areas.

Even without this the database can be used in both qualitatively and quantitatively to support assessments of the “policy impact” that a body of research has had.

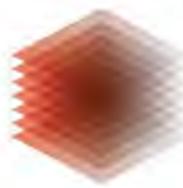
One use is discovery: many institutions and funders are not aware of where the research they support is cited in policy, or if it is at all. Overton allows these users to quickly identify where they might want to focus their attention when looking for candidates for further, qualitative analysis and case studies.

Another is more quantitative: enough scholarly articles are cited in policy and picked up by Overton to allow for “like for like” queries that can control for article age and subject area.

In simple terms this involves checking a large list of “target” articles against the set cited in policy, and then comparing it to a number of “comparand” control sets made of up articles published in the same journal and in a similar timeframe. If the %age of target articles cited in policy is higher than the average %age of comparand articles cited in policy then one can make a case that they are being cited more than expected: though of course further analysis is required to determine exactly why.

## References

- Altmetric. (2021). *Policy Documents*. Retrieved from help.altmetric.com: <https://help.altmetric.com/support/solutions/articles/6000236695-policy-documents>
- Baron, J. (2018). A Brief History of Evidence-Based Policy. *The ANNALS of the American Academy of Political and Social Science*.
- Defining impact*. (2022). Retrieved from UK Research and Innovation: <https://www.ukri.org/councils/esrc/impact-toolkit-for-economic-and-social-sciences/defining-impact/>
- European Parliamentary Research Service. (2021). *Evidence for policy-making: Foresight-based scientific advice*.
- Intergovernmental Panel of Climate Change. (2013). *IPCC Factsheet: What literature does the IPCC assess?* IPCC.
- MacKillop, E. (2019). *Does knowledge brokering facilitate evidence-based policy? A review of existing knowledge and an agenda for future research*. Policy Press.
- Newson, R. (2018). Does citation matter? Research citation in policy documents as an indicator of research impact – an Australian obesity policy case-study. *Health Research Policy and Systems*.
- Overton. (2023). *Indexed policy sources*. Retrieved from app.overton.io: [https://app.overton.io/policy\\_sources.php](https://app.overton.io/policy_sources.php)
- Patience, G. (2017). Citation analysis of scientific categories. *Heliyon*.
- Scopus. (2023). *What are Scopus subject area categories and ASJC codes?* Retrieved from Scopus Support Centre: [https://service.elsevier.com/app/answers/detail/a\\_id/12007/supporthub/scopus/related/1/](https://service.elsevier.com/app/answers/detail/a_id/12007/supporthub/scopus/related/1/)
- Smith, K. (2020). *The Impact Agenda: Controversies, Consequences and Challenges*. Policy Press.
- Szomszor, M. (2022). Overton: A bibliometric database of policy document citations. *Quantitative Science Studies*.
- Wellcome Trust. (2020). *Introducing Reach: find and track research being put into action*. Retrieved from Wellcome Data: <https://medium.com/wellcome-data/introducing-reach-find-and-track-research-being-put-into-action-dec2a2fca93b>



**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](http://WWW.TIB.EU)

## ICSTI Member Survey: Infrastructures and Services for non-textual materials

**Margret Plank**, TIB - German National Library for Science and Technology, Germany  
**Lisa Curtin** and **Shelby Stooksbury**, DOE Office of Scientific and Technical Information (OSTI), United States

### Video Presentation

<https://av.tib.eu/media/59861>

### Introduction

The challenges posed by changes in research and publishing behavior, by the upheaval of the publication market and by the emergence of very different forms of distribution and use, have evolved and continue to change the tasks in information supply. The increasing digitization of workflows and processes is having a significant influence on the demands that customers and users place on modern information services (TIB 2027). These ongoing dynamic developments in the area of information supply and knowledge development are influencing and changing science and technology organizations. Today, science and technology organizations function as modern information infrastructure facilities that claim to provide researchers with innovative services throughout the entire scientific process. In order to meet the user's expectations, they must continuously analyze the changing needs of the scientific community, monitor trends, implement technical innovations and develop services tailored to the needs of their target groups. The International Council for Scientific and Technical Information (ICSTI) founded in 1984, is a not-for-profit membership organization offering a forum for interaction between national and international organizations that create, disseminate, and use scientific and technical information. In order to provide a more accurate and multifaceted picture of best practices regarding innovative services among ICSTI members, we conducted an online survey. The focus of the survey was on the provision of infrastructures, tools and services for the publication, use and creation of non-textual science and technology materials such as audiovisual media, 3D objects, graphics, research data and research software which have been little investigated in the science, technology and information context. The present study aims to contribute to a better understanding of what ICSTI member organizations around the globe already offer their scientific user groups regarding non-textual materials and what is planned for the future. Good practices serve as examples for other organizations. With this in mind, we hope that the results of the study will be of benefit for other information centers and technical libraries.

### 2. Study Design

The questionnaire used for the survey mostly consisted of closed-ended questions. The assigned method was quantitative (questionnaire) and qualitative (free-text fields). The survey included questions about the provision of research tools, infrastructures and services for the publication, access, use and creation of non-textual materials. Some questions also aimed at the provision of training with regard to non-textual publications and the assignment of Digital Object Identifiers (DOI). The survey was sent to representatives of all nineteen full ICSTI member organizations, namely: the Agricultural Information Institute of the Chinese Academy of Agricultural Sciences (All of CAAS), ETH-Bibliothek, the German National Library of Science & Technology (TIB), the International Atomic Energy Agency (IAEA), the Institute of Medical Information of the Chinese Academy of Medical Sciences (IMI CAMS), the Institute of Scientific and Technical Information of China (ISTIC), the International Federation of Library Associations (IFLA),

the International Standard Serial Number (ISSN) International Centre, the Japan Science and Technology Agency (JST), King Abdullah University of Science and Technology (KAUST), the Korea Institute of Science and Technology Information (KISTI), the National Research Council of Canada (NRC), the National Science and Technology Library (NSTL) of China, the National Science Library of the Chinese Academy of Sciences (NSLC), the Shanghai Society for Scientific and Technical Information (SSSTI), the State and University Library of Gottingen (SUB), the U.S. Department of Energy Office of Scientific and Technical Information (OSTI), the U.S. National Library of Medicine, and VTT Technical Research Centre of Finland. These organizations were given a duration of two months (08/22/22-10/19/22) to complete the survey.

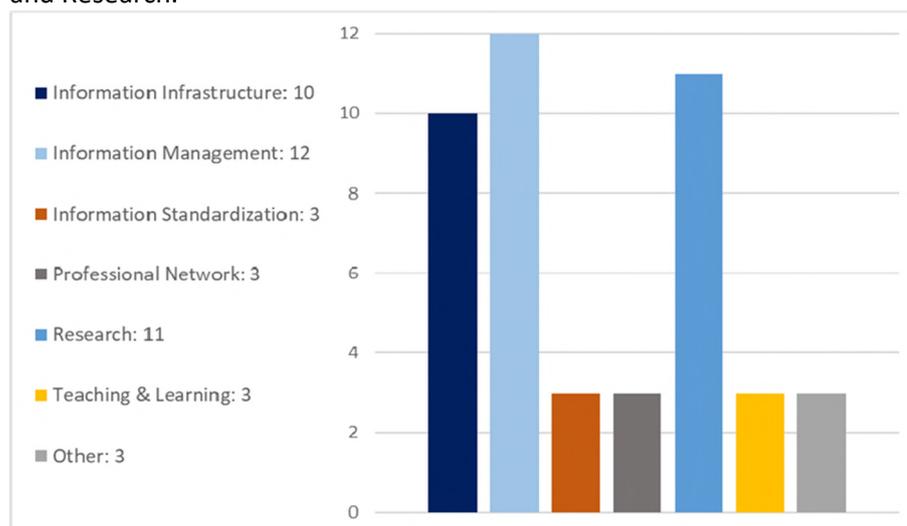
The total number of participant organizations that completed the survey was fifteen. Due to the small sample size, the results of this survey do not claim to be representative.

### 3. Results

#### 3.1 Formal Criteria

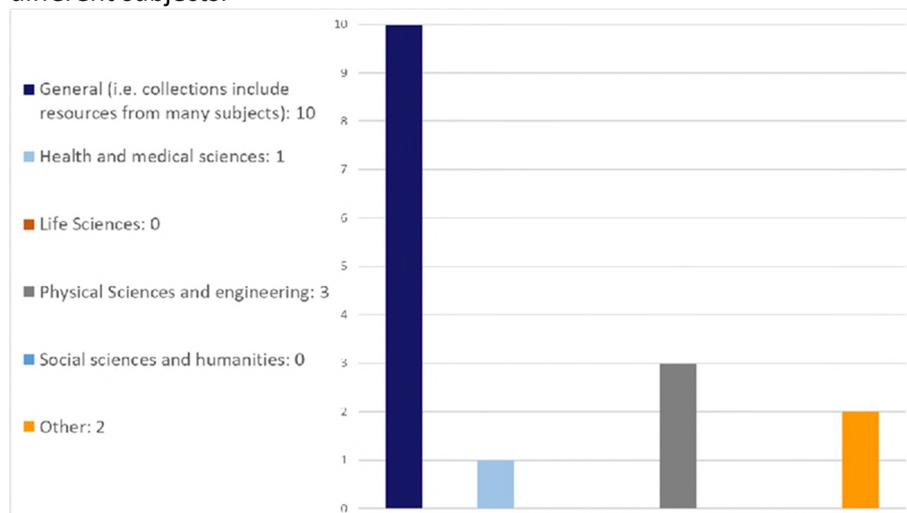
##### 3.1.1 What are your organization’s major functions or activities?

The most frequent responses were Information Management, Information Infrastructure and Research.



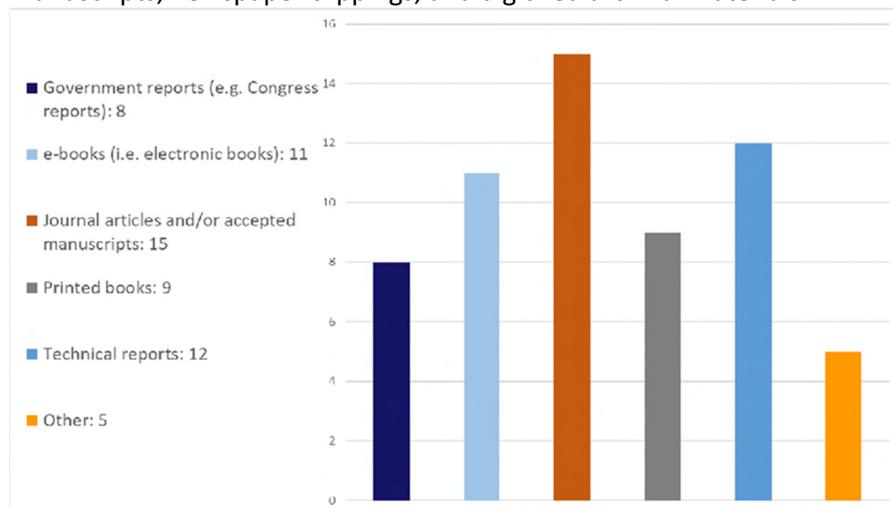
##### 3.1.2 What is the primary subject focus of your organization?

Most of the respondents answered that their collections contained resources from many different subjects.



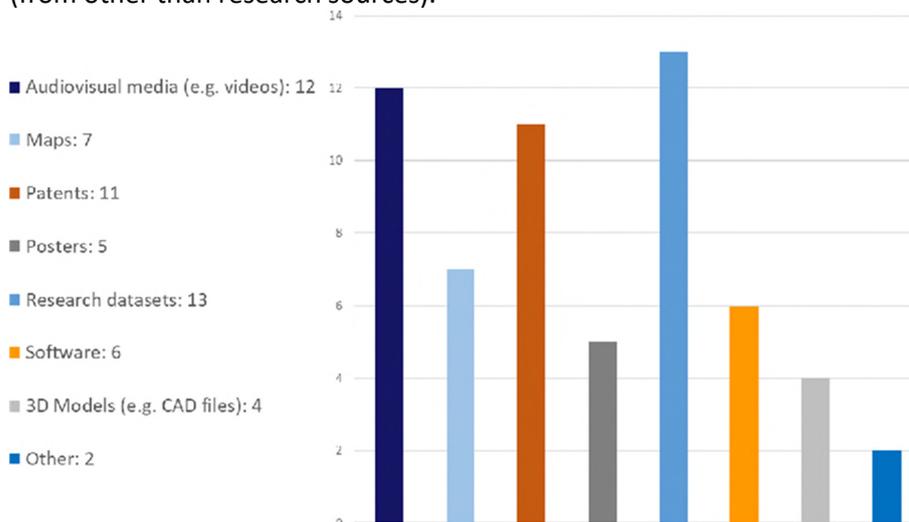
**3.1.3 What types of digital textual scientific and technical information materials does your organization collect or store?**

The collected textual formats included journal articles, technical reports, e-books, government reports and others such as theses/dissertations, program documents, printed books, clinical vocabularies and value sets, lab notes and notebooks, handwritten manuscripts, newspaper clippings, and digitized archival materials.



**3.1.4 What types of digital non-textual scientific and technical information materials does your organization collect or store?**

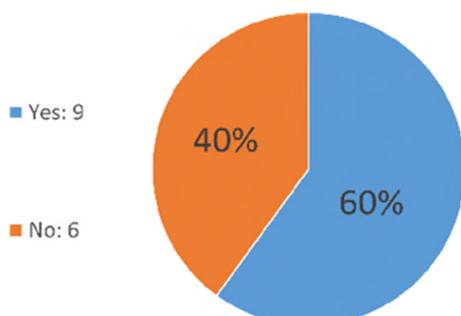
A wide range of object types were named. In the order of most frequently mentioned: research datasets (such as measurement data, material samples, structural formulas, simulation data, and geospatial data), audiovisual material, patents, maps, software, posters, 3D models (CAD) and others such as oral histories, prints & photographs, vocabularies and related health data standards, and born-digital publications and datasets (from other than research sources).



### 3.2 Audiovisual Media

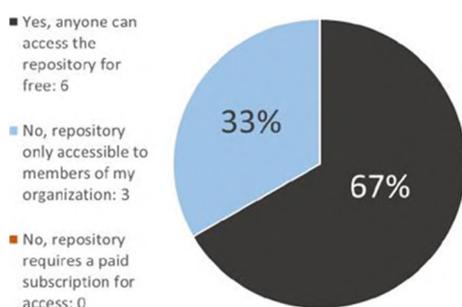
#### 3.2.1 Does your organization host a repository for audiovisual media?

Nine respondents answered yes, with six answering no. From the nine who answered yes, seven said the repository was exclusive to containing audiovisual media.



#### 3.2.2 Is the audiovisual media repository publicly accessible?

Six respondents answered that the repository was publicly available, whereas three answered that the repository was only accessible internally to members of that organization.

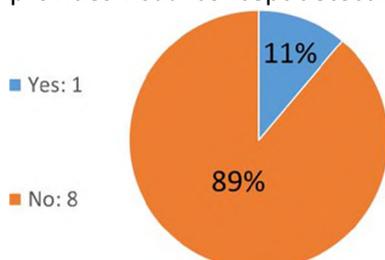


#### 3.2.3 If your organization does not host a repository for audiovisual media, what, if any, external services are used to provide access?

The named external services used to provide access to audiovisual media were Youtube, TikTok, Youku, [WorldWideScience.org](http://WorldWideScience.org), J-Stage Data and NLM Digital Collections. Additionally, a collection of DVDs that can be accessed locally was mentioned.

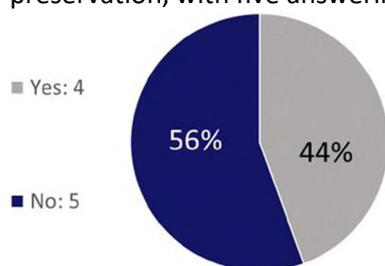
#### 3.2.4 Does your organization provide a visual search tool for the audiovisual media repository?

Only one respondent answered that their organization uses a visual search tool that provides visual concept detection based on neural networks and deep learning methods.



**3.2.5 Does your organization provide long-term digital preservation for audiovisual media?**

Four respondents answered that their organization provides long-term digital preservation, with five answering that they do not.

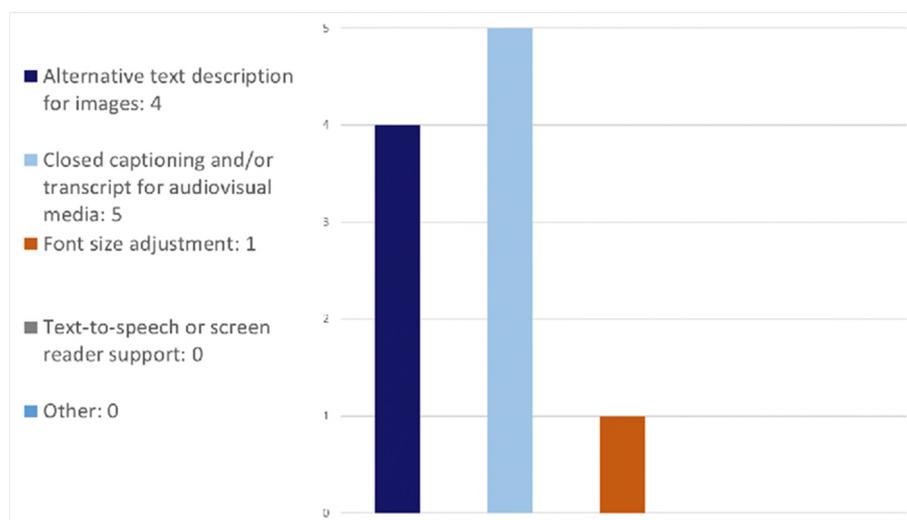


**3.2.6 Does your organization provide any of the following infrastructures, tools, or services to make the audiovisual media repository accessible to users with disabilities?**

Regarding the provision of tools and infrastructures for users with disabilities, the organizations provide services such as closed captioning or transcripts, alternative text description and text size adjustments for audiovisual material.

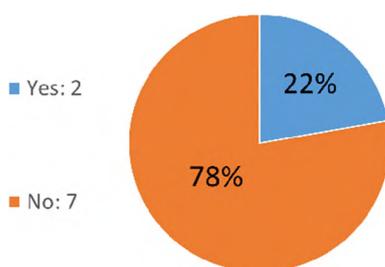
**3.2.7 Does your organization provide any of the following training resources in relation to the management or publication of audiovisual media?**

Most of the organizations provide written text or guidance, as well as instructional videos used as training resources in relation to management or publication of audiovisual media.



**3.2.8 Does your organization assign persistent identifiers to audiovisual media?**

Only two organizations answered that they assign persistent Identifiers for audiovisual media, those being specifically Digital Object Identifiers (DOIs).



**3.2.9 Which DOI registration services are used to assign DOIs to audiovisual media?**

In addition to DataCite’s DOI registration service, one organization utilizes the Entertainment Identifier Registry (EIDR).

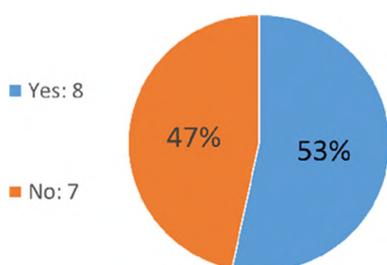
**3.2.10 Does your organization assign any of the following types of persistent identifiers to audiovisual media (in addition to or instead of DOIs)?**

None of the organizations use other persistent identifiers such as Arxiv or URN.

**3.3 Images & Graphics**

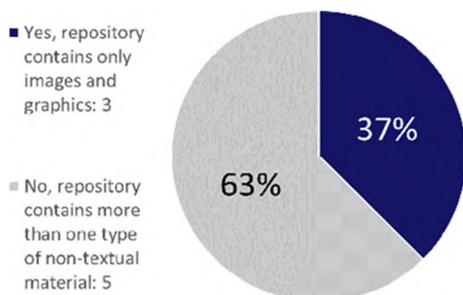
**3.3.1 Does your organization host a repository for Images and Graphics?**

Eight of the respondents answered that their organization had a repository for images and graphics, while the other seven stated that their organization did not.



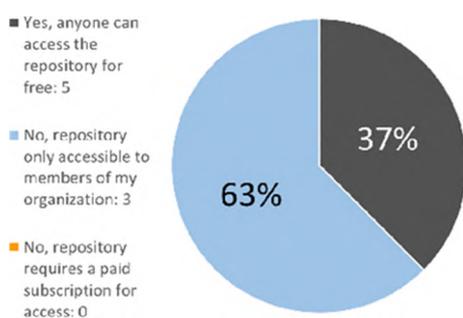
**3.3.2 Is the repository exclusive to Images and Graphics?**

Five respondents claimed that their repository was not exclusive to images and graphics, and three replied that their repository was exclusive for images and graphics.



**3.3.3 Is the repository for Images and Graphics publicly accessible?**

Five organizations stated that their repository was publicly available, while three answered that the repository was only accessible to members of their organization.

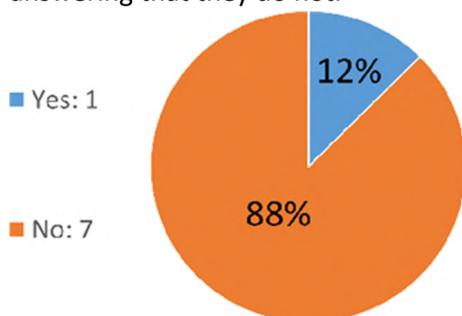


### 3.3.4 If your organization does not host a repository for images and graphics, what, if any, external services are used to provide access?

The named external services used to provide access to images and graphics were Youtube, TikTok, Youku, J-Staged, Figshare, Zenodo and Wikicommons. Additionally, a collection of analogue maps that can be accessed locally was listed.

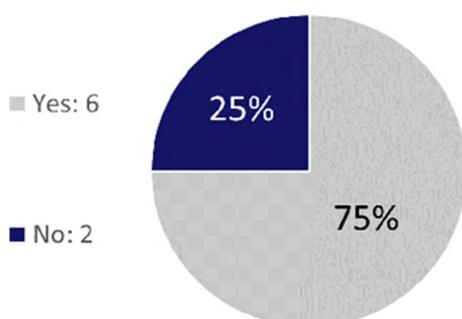
### 3.3.5 Does your organization provide a visual search tool for the images and graphics repository?

Only one respondent replied that their organization had a visual search tool, with the rest answering that they do not.



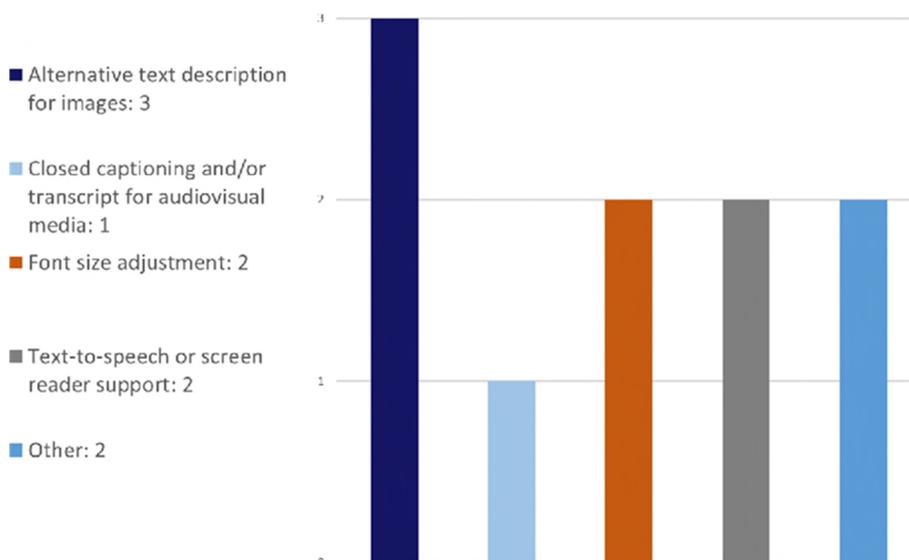
### 3.3.6 Does your organization provide long-term digital preservation for images and graphics?

Six of the respondents answered that their organization provides long-term digital preservation for images and graphics, while the other two stated that their organization did not.



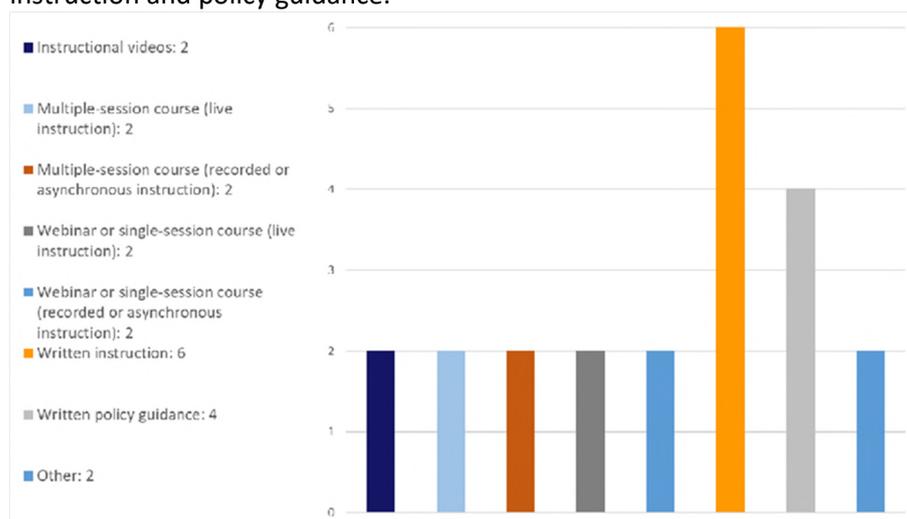
### 3.3.7 Does your organization provide any of the following infrastructures, tools, or services to make the images and graphics repository accessible to users with disabilities?

Regarding the provision of tools and infrastructures for users with disabilities, the organizations provide services such as alternative text description, text size adjustments, text-to-speech or screen reader support and closed captioning and/or transcripts for images and graphics.



**3.3.8 Does your organization provide any of the following training resources in relation to the management or publication of images and graphics?**

With respect to training materials, the most prevalent resource provided is written instruction and policy guidance.



**3.3.9 Does your organization assign persistent identifiers to images and graphics?**

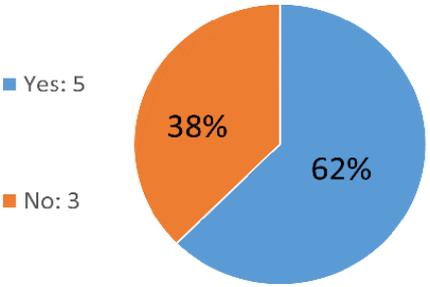
Five organizations assign persistent identifiers to images and graphics.

**3.3.10 Does your organization assign digital object identifiers (DOIs) to images and graphics? Which DOI registration services are used to assign DOIs to images and graphics?**

Three organizations assign DOIs to images and graphics using either the DataCite or KISTI registration services.

**3.3.11 Does your organization assign any of the following types of persistent identifiers to images and graphics (in addition to or instead of DOIs)?**

The responses included internal persistent identifiers assigned to images and graphics.

<p>Does your organization assign persistent identifiers to images and graphics?</p>	<p>Does your organization assign digital object identifiers to images and graphics?</p> <ul style="list-style-type: none"> <li>• Yes: 3</li> <li>• No: 2</li> </ul>
 <p>■ Yes: 5 62%</p> <p>■ No: 3 38%</p>	<p>Which DOI registration services are used to assign DOIs to images and graphics?</p> <ul style="list-style-type: none"> <li>• DataCite: 3</li> <li>• Entertainment Identifier Registry (EIDR): 0</li> <li>• ISTIC &amp; Wanfang Data: 0</li> <li>• Japan Link Center (JaLC): 0</li> <li>• Korea Institute of Science and Technology (KISTI): 1</li> <li>• Multilingual European DOI Registration Agency: 0</li> <li>• Publications Office of the European Union: 0</li> <li>• Other: 0</li> </ul> <p>Does your organization assign any of the following types of persistent identifiers to images and graphics (in addition to or instead of DOIs)?</p> <ul style="list-style-type: none"> <li>• ARK (archival resource key): 0</li> <li>• ArXiv: 0</li> <li>• Internal identifier (i.e. organization specific identifier): 1</li> <li>• URN (uniform resource name): 0</li> <li>• Other: 1</li> </ul>

### 3.4 Research Datasets

#### 3.4.1 Does your organization host a repository for research datasets?

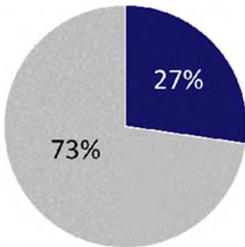
Eleven organizations answered that they do host a repository for research datasets, while four do not.

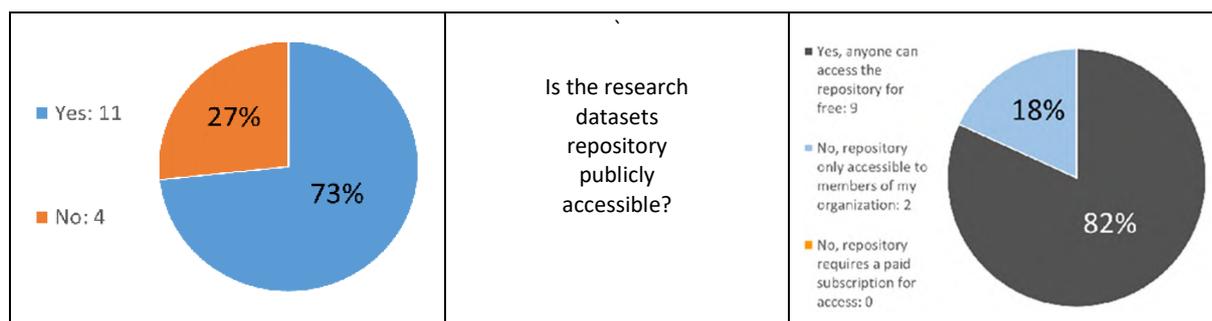
#### 3.4.2 Is the repository exclusive to research datasets?

Eight out of the eleven organizations stated that their repository was not exclusive to research datasets; three of the organizations' repositories only contain research datasets.

#### 3.4.3 Is the research datasets repository publicly accessible?

Nine respondents replied that their repository for research datasets is open to everyone, while two answered that it was only accessible within their organization.

<p>Does your organization host a repository for research datasets?</p>	<p>Is the repository exclusive to research datasets?</p>	<ul style="list-style-type: none"> <li>■ Yes, repository contains only research datasets: 3</li> <li>■ No, repository contains more than one type of non-textual material: 8</li> </ul>  <p>27%</p> <p>73%</p>
--	--	---

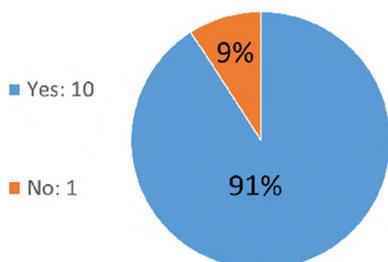


**3.4.4 If your organization does not host a repository for research datasets, what, if any, external services are used to provide access?**

The named external services used to provide access to research data were Zenodo, GitHub and Radar-Service.

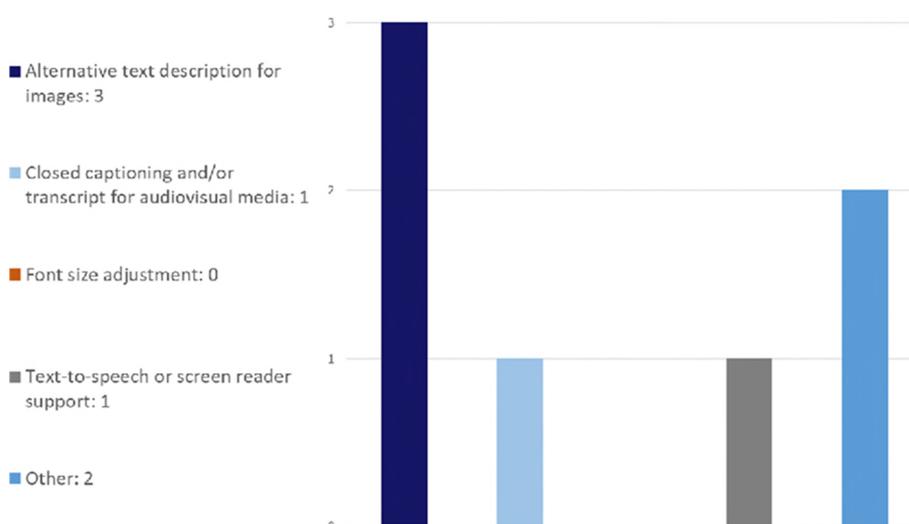
**3.4.5 Does your organization provide long-term digital preservation for research datasets?**

Ten organizations stated they provide long-term digital preservation for their datasets, while one organization does not.



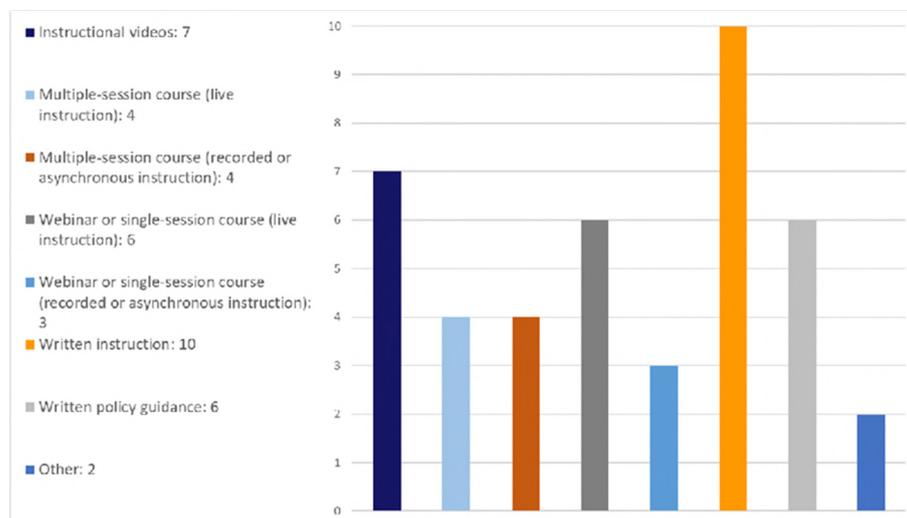
**3.4.6 Does your organization provide any of the following infrastructures, tools, or services to make the research datasets repository accessible to users with disabilities?**

Regarding the provision of tools and infrastructures for users with disabilities, the organizations provide services such as alternative text description, text-to-speech or screen reader support and closed captioning and/or transcripts for research datasets.



**3.4.7 Does your organization provide any of the following training resources in relation to the management or publication of research datasets?**

The responses concerning training material included written instruction and policy guidance, instructional videos, webinars and multiple-session courses either live or recorded.



**3.4.8 Does your organization assign persistent identifiers to research datasets?**

Nine organizations answered that they assign persistent identifiers to the datasets, while two of them do not.

**3.4.9 Does your organization assign digital object identifiers (DOIs) to research datasets? Which DOI registration services are used to assign DOIs to research datasets?**

Seven respondents declared their organizations assign DOIs to research datasets, most using DataCite. Other registration services used are ISTIC & Wanfang Data, JaLC and KISTI.

**3.4.10 Does your organization assign any of the following types of persistent identifiers to research datasets (in addition to or instead of DOIs)?**

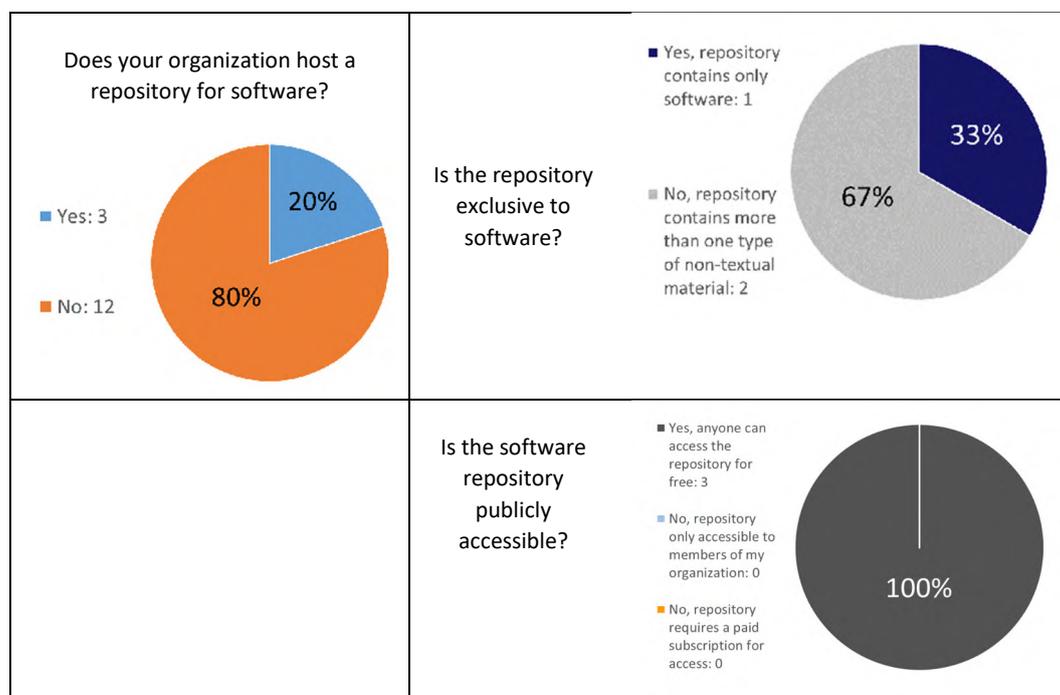
The responses included the use of internal Identifiers and other identifiers such as Permalink and EPIC handles.

<p>Does your organization assign persistent identifiers to research datasets?</p> <p>■ Yes: 9 ■ No: 2</p>	<p>Does your organization assign digital object identifiers to research datasets?</p> <ul style="list-style-type: none"> <li>• Yes: 7</li> <li>• No: 2</li> </ul> <p>Which DOI registration services are used to assign DOIs to research datasets?</p> <ul style="list-style-type: none"> <li>• DataCite: 5</li> <li>• Entertainment Identifier Registry (EIDR): 0</li> <li>• ISTIC &amp; Wanfang Data: 1</li> <li>• Japan Link Center (JaLC): 1</li> <li>• Korea Institute of Science and Technology (KISTI): 1</li> <li>• Multilingual European DOI Registration Agency: 0</li> <li>• Publications Office of the European Union: 0</li> <li>• Other: 0</li> </ul> <p>Does your organization assign any of the following types of persistent identifiers to research datasets (in addition to or instead of DOIs)?</p> <ul style="list-style-type: none"> <li>• ARK (archival resource key): 0</li> <li>• arXiv: 0</li> <li>• Internal identifier (i.e. organization specific identifier): 3</li> <li>• URN (uniform resource name): 1</li> <li>• Other: 2</li> </ul>
---	--

### 3.5 Software

#### 3.5.1 Does your organization host a repository for software? Is the repository exclusive to software? Is the software repository publicly accessible?

Three of the respondents stated that they have a repository for software, with only one of those being exclusive to containing software only. All three software repositories are publicly available.



#### 3.5.2 If your organization does not host a repository for software, what, if any, external services are used to provide access?

The named external services used to provide access to software were GitHub, Zenodo and Microsoft Teams.

#### 3.5.3 Does your organization provide long-term digital preservation for software?

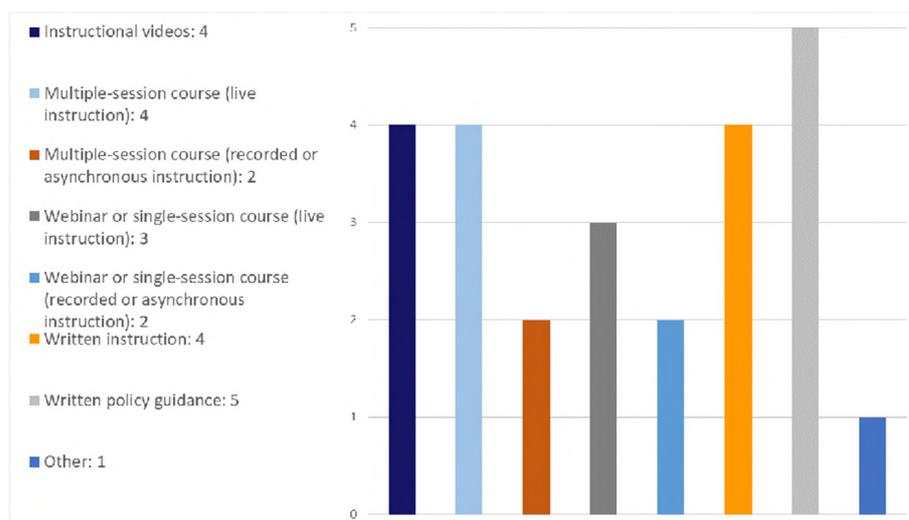
All three of the organizations who host a repository for software provide long-term preservation.

#### 3.5.4 Does your organization provide any of the following infrastructures, tools, or services to make the software repository accessible to users with disabilities?

None of the organizations provide specific tools or services for software for users with disabilities.

#### 3.5.5 Does your organization provide any of the following training resources in relation to the management or publication of software?

Training resources provided include written instruction and policy guidance, webinars and multiple-session courses either live or recorded, and instructional videos.



**3.5.6 Does your organization assign persistent identifiers to software? Does your organization assign digital object identifiers (DOIs) to software? Which DOI registration services are used to assign DOIs to software?**

All three organizations which provide a repository for software assign persistent identifiers in the form of DOIs using the DataCite registration service.

**3.5.7 Does your organization assign any of the following types of persistent identifiers to software (in addition to or instead of DOIs)?**

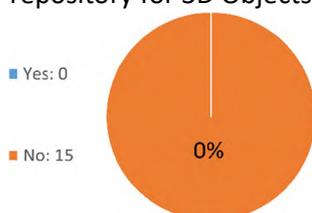
One organization assigns internal identifiers to software.

<p>Does your organization assign persistent identifiers to software?</p> <p>■ Yes: 3 ■ No: 0</p> <p>100%</p>	<p>Does your organization assign digital object identifiers to software?</p> <ul style="list-style-type: none"> <li>• Yes: 3</li> <li>• No: 0</li> </ul> <p>Which DOI registration services are used to assign DOIs to software?</p> <ul style="list-style-type: none"> <li>• DataCite: 3</li> <li>• Entertainment Identifier Registry (EIDR): 0</li> <li>• ISTIC &amp; Wanfang Data: 0</li> <li>• Japan Link Center (JaLC): 0</li> <li>• Korea Institute of Science and Technology (KISTI): 0</li> <li>• Multilingual European DOI Registration Agency: 0</li> <li>• Publications Office of the European Union: 0</li> <li>• Other: 0</li> </ul> <p>Does your organization assign any of the following types of persistent identifiers to software (in addition to or instead of DOIs)?</p> <ul style="list-style-type: none"> <li>• ARK (archival resource key): 0</li> <li>• arXiv: 0</li> <li>• Internal identifier (i.e. organization specific identifier): 1</li> <li>• URN (uniform resource name): 0</li> <li>• Other: 0</li> </ul>
--	---

### 3.6 3D Objects

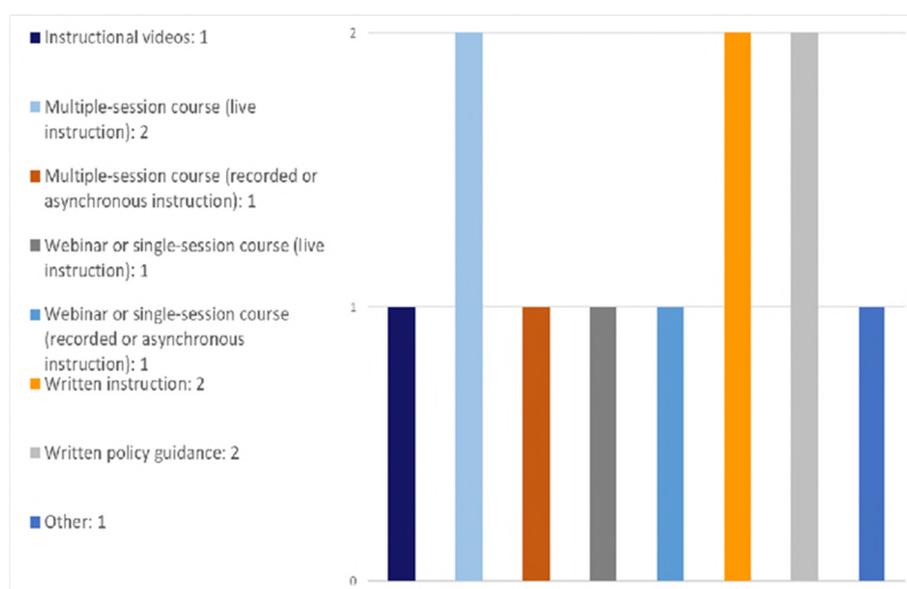
#### 3.6.1 Does your organization host a repository for 3D objects?

Fifteen out of the fifteen participating organizations declared that they do not provide a repository for 3D Objects.



#### 3.6.2 Does your organization provide any of the following training resources in relation to the management or publication of 3D objects?

Although no repositories are available for 3D objects within the organizations surveyed, some of them provide training resources such as written instruction or policy guidance, instructional videos, and webinars or multiple session courses (either live or recorded) for their users.



### 3.7 Challenges

#### 3.7.1 What are the biggest challenges when it comes to providing infrastructures, tools and services for non-textual materials within your organizations?

The majority of the organizations surveyed indicated that they would like to provide more services in the area of non-textual materials, but the effort to meet the demands of the different types of objects was simply too complex and time consuming. According to the responses other challenges these organizations face include:

- Definition of user requirements
- Construction of the ecosystem and cooperation mechanism, including the contributors, users, funders, administrators, and commercial partners
- Communication with researchers about the expectations for submitting non-textual materials to their organization
- Long-term preservation of the materials
- Provision of resources to support staff, IT infrastructure, and financial resources
- Accessibility, findability, semantic relations and visualizations of relations between data, text graphics, software and audiovisual media.
- Clarification of copyright for videos

#### 4. Summary and Conclusion

The survey conducted primarily provided an indication of existing tools, infrastructures and services of the participating ICSTI member organizations regarding non-textual materials. Non-textual material such as research data, research software, graphics, 3D Models and audiovisual media is research output and should thus be accessible and reusable. This concerns first and foremost the data on which the research results are based. Beyond that, however, the software used to generate or analyze the data is just as important. In addition, there are visualizations such as graphics, 3D models, animations, simulations or videos that enable understanding and interpreting of the data.

It appears that research data infrastructures and management are already well established. 75% of the respondents stated that their organization operates its own repository, most of which are publicly available. The vast majority of datasets have a DOI and are archived for the long-term. Nevertheless, there is still a need for improvement, one example being the reusability of research data. At the same time, the study showed that comparable activities in the field of scientific software, i.e. the management of research software, have hardly been established. Only three organizations stated they run their own repository for software, from which only one is publicly available. However some organizations are using GitHub and Zenodo to store the software. In addition, media-specific solutions are required for other non-textual materials, such as audiovisual media, graphics or 3D objects, which according to our study seem to hardly exist in the surveyed organizations. In terms of audiovisual media, just over half of the surveyed organizations have their own repository. Alternatives used are commercial platforms, such as YouTube. Only one organization uses a visual search tool for better discovery in videos. Only two organizations assign DOIs to audiovisual media. For 3D objects, none of the organizations surveyed has a solution. The survey brought to light that most non-textual materials are not yet accessible and citable in the long term.

The FAIR Data Principles were published by the FORCE11 Group in 2016 (Wilkinson 2016) and can be understood as a guideline, where FAIR stands for Findable, Accessible, Interoperable and Reusable. Even though the FAIR principles were inspired by Open Science, these two movements are explicitly not to be equated (Mons 2017). However, the FAIR principles require "clarity and transparency about the conditions for access and re-use" of the data, even if they are not "open" (Mons 2017). In principle, this means that organizations which have a focus on Information Management, Infrastructures and Research, as the organizations surveyed in this study do, should provide corresponding infrastructures and services which support the FAIRness of their data. These services should be developed in close cooperation with the users, so that any research output, whether textual or non-textual will be accessible and re-usable in the future and thus contributing to the comprehensibility of research.

#### 5. Literature

Mons, B., Neylon, C., Velterop, J., Dumontier, M., Da Silva Santos, L. and Wilkinson, M. (2017) 'Cloudy, increasingly FAIR; Revisiting the FAIR Data guiding principles for the European Open Science Cloud', *Information Services and Use*. 37 (1): pp. 49-56. doi: 10.3233/ISU-170824

Technische Informationsbibliothek 2017. TIB Survey 2017 on information procurement and publishing behavior of researchers in the natural sciences and engineering: Technische Informationsbibliothek (TIB) DOI <https://doi.org/10.34713/apm4-rm61>

Wilkinson, M.D., et al. (2016) The fair guiding principles for scientific data management and stewardship. *Sci. Data* 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

# Czech National Repository of Grey Literature

The logo for NUSL (National Library of Technology) features the letters 'NUSL' in a white, sans-serif font. The letter 'U' is stylized with a vertical bar on its right side. The logo is set against a dark grey square background, which is itself on a light green background.

## NUSL is

a digital  
repository  
for grey  
literature

## Free

online  
access

## Features

### Provider:

National Library of Technology  
Prague, Czech Republic

### Records:

over 500,000 records

### Collection provenance:

Czech Republic

### Partners:

over 150 organizations (Academy of Science,  
Public Research Institutions, Universities, State  
Offices, Libraries, NGOs etc.)

### International Cooperation:

OpenGrey, OpenAire, ROAR, OpenDOAR, BASE,  
WorldWideScience

## Goals

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

## What else?

Conference on Grey Literature and  
Repositories

<https://nusl.techlib.cz/en/conference>

Informative Webpages

<https://nusl.techlib.cz/en/>

[www.nusl.cz](http://www.nusl.cz)

**NTK**  
25° 90° 105°  
Národní technická knihovna  
National Library of Technology

**NUSL**  
czech  
national  
repository  
of grey  
literature

## Preregistration of research for theses - a new standard?

Tereza Šimová, Czech University of Life Sciences Prague, Czech Republic

### Video Presentation

<https://av.tib.eu/media/58579>

#### **Abstract:**

*The reproducibility crisis is the ongoing methodological crisis, where it is difficult or impossible to reproduce the results of many scientific studies. Because the reproducibility of scientific results is an essential part of science, such failures undermine the credibility of the theories and potentially call into question a substantial part of scientific knowledge. There are several ways to strengthen the reproducibility of results, such as proper reporting and dissemination, e.g., pre-registration of research, registered reports, open data, open code, open research notes, or open access. In this paper, I propose incorporating preregistration into the theses writing process for undergraduate students, which should lead to enhanced reproducibility, transparency, and openness of research results in the context of final theses. Pre-registration is a data-marked, predetermined, read-only research roadmap that, in most cases, is created before the study and published in a public repository. Preregistration increases the credibility, transparency, openness, and robustness of students' results in writing their theses. It also serves as a tremendous educational and communicative element of the thesis writing process that raises the level of student theses and demonstrates a certain level of student knowledge and skills. In this paper, I propose a timeframe for incorporating pre-registration into theses writing processes and discuss the challenges surrounding this process from the perspective of key stakeholders - policymakers, universities, theses supervisors, and students. The article also presents practical advice on where to find support and further guidance on pre-registration aspects. This paper thus pioneers the idea of incorporating preregistration into the normal thesis writing process.*

*Keywords: reproducibility, undergraduate, students final work, open science*

#### **Replication crisis**

The reproducibility crisis is the ongoing methodological crisis, where it is difficult or impossible to reproduce the results of many scientific studies. Because the reproducibility of scientific results is an essential part of science, such failures undermine the credibility of the theories and potentially call into question a substantial part of scientific knowledge (Nosek *et al.*, 2015, 2018; Munafò *et al.*, 2017; Alger, 2020; Grahe, 2021). Many things contribute to this, but in general, many of these issues come down to this hypothetical deductive model of science being short-circuited by questionable research practices (Center for Open Science, no date). Lack of replication, Low statistical power (p-hacking), HARKing, Publication bias, and lack of data sharing all of these and other questionable research practices have contributed to creating a reproducibility crisis (Quintana, 2015; Munafò *et al.*, 2017; Center for Open Science, no date).

There are several ways to strengthen the reproducibility of results, such as proper reporting and dissemination, e.g., pre-registration of research, registered reports, open data, open code, open research notes, or open access. Nosek *et al.* (2018) argue that in an ideal world, the basics of preregistration are taught in primary school. I do not go that far in my proposal. In my proposal, I discuss the possibility of incorporating preregistration into the theses writing process for undergraduate students, which should lead to enhanced reproducibility, transparency, and openness of research results in the context of final theses. I propose a timeframe for incorporating pre-registration into theses writing processes and discuss the challenges surrounding this process from the perspective of key stakeholders - policymakers, universities, theses supervisors, and students.

**Preregistration**

Pre-registration is documenting a research plan, which usually occurs before data collection and analysis. As Nosek et al. (2018) point out, preregistration for research is a commitment to analytical steps without prior knowledge of the research results. Grahe (2021, p. 33) define preregistration as an "effort to freeze, or predetermine, those decisions before the data are collected by requiring the researcher to document the decisions and then date-stamp them using a permanent repository." Pre-registration can also be thought of as a research roadmap - a data-marked, predetermined, read-only research plan that in most cases is created before the study itself and published in a public repository. Pre-registration, in its simplest form, may involve only the registration of the primary study design but it is possible to create very detailed pre-specification of the study procedures, preliminary outcomes, and statistical analysis plan (Munafò *et al.*, 2017).

There are different types of preregistrations. For example, unreviewed preregistration, reviewed preregistration (Registered Reports), or Registered Replication Reports. However, this paper does not aim to present comprehensive information about these preregistrations. However, I think it is important to mention that when it is written about preregistration in this paper, I am referring primarily to classical unreviewed preregistrations.

**Why should preregistration matter during theses writing?**

Preregistration brings several benefits to research, which are discussed in several publications - see e.g. (Quintana, 2015; Munafò *et al.*, 2017; Cockburn, Gutwin and Dix, 2018; Nosek *et al.*, 2018, 2019; Lakens, 2019; Grahe, 2021). It is, therefore, not surprising that the implementation of preregistration as part of theses writing would also bring some benefits. The advantage of preregistration in both research and the proposal here is the increased credibility, transparency, and robustness of the results. As Button (2018) wrote, final theses often become an exercise in compounding the research design limitations, which can be frustrating for both the student and the supervisor. By having students preregister their research, this frustration could be alleviated. Another advantage of preregistration is its educational benefit. By getting students interested in data and analysis at an early stage of research, their knowledge of data analysis and management, as well as research practices (such as reproducibility, transparency, and open science principles), will be enhanced. Students will think about their research design right at the beginning of the thesis writing, allowing them to avoid mistakes in the research approach. Button (2018) argues that early training will bring comfort and creativity to students concerning similar approaches later in their research careers. Grahe (2021) shares this argument, describing preregistration as a way to "front-loading the writing process and reducing challenges of miscommunication between collaborators during the collection and analysis step." Preregistering can also help students for their future selves because sometimes we just forget what we planned. Preregistration can remind students what they planned, what data and analyses they wanted to use, what their sample was supposed to be, etc. This sort of reminder also reduces self-deception, which is another benefit of preregistration. The advantage is also easier communication of research, which is sometimes difficult for students. Preregistration give us and to our colleagues a clear roadmap to our research ideas, hypotheses, data collection, analysis etc. In certain cases, pre-registering research can help students to gain a greater status (or recognition) from the scientific community. By preparing a preregistration, students demonstrate that they have a certain level of skills and knowledge that will be useful not only in their theses defence, but also in their possible future careers (not only in research related jobs).

Sometimes people cheat, and no one can stop it. The problem, as Grahe (2021, p. 36) notes, is that "in science, we can cheat without knowing we are doing it. It all comes down to the reliability of the probability that a statistic is unlikely to have occurred by chance, the p-value". Pre-registration helps students to recall each of the decisions before the

data appear. The student can then more clearly understand the difference between an exploratory question, for which a p-value adjustment must be made, and confirmatory research, for which it is not.

### **Pioneers in the implementation of preregistration in theses writing**

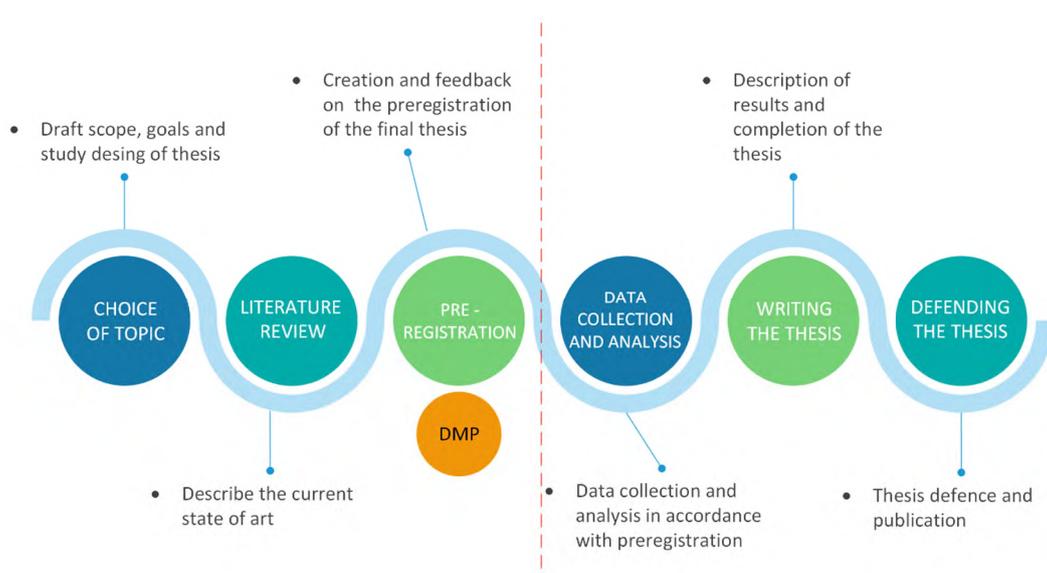
A number of academic researchers has already incorporated the use of preregistration and other transparent research methods at various universities, such as the Collaborative Replication and Education Project (CREP) and the Psychological Science Accelerator course (Crchartier, 2018) or the Reproducibility Education in an Undergraduate Capstone Course (*Reproducibility Education in an Undergraduate Capstone Course*, 2021), or the GW4 Undergraduate Psychology Consortium (Button, 2018).

At the same time, it is clear that this approach is not suitable for all types of research. As Button (2018) suggests, in certain types of laboratories, where laboratory equipment is expensive and/or the layout of the laboratory does not allow for standardized operating procedures, it is not so easy to pre-register. For these topics where preregistration does not make sense, it may be possible to consider replacing preregistration with a data management plan.

### **Timeframe for incorporating preregistration into the theses writing process**

As Grahe (2021) wrote, there is very little time for students during the process of theses writing. Thus is essential to think about the right time for incorporating preregistration into the theses writing process. Figure 1 shows my suggestion of how preregistration can be incorporated into the timeframe of writing a theses. In the circle can be seen the time period, and the text above or below the circle describes the goal of a particular period. What is important is that the students receive the first feedback on their research design in the early parts of the research (the red dashed line). This early feedback will help students avoid potential mistakes in the research approach and also will help them produce higher-quality theses. In this way, students will not only learn to do good research practice, but more importantly, they will become aware of their research and what they should look out for. Also, students will receive two pieces of feedback from the scientific community during the writing process (the first during pre-registration and the second during the defense).

Figure 1 Incorporating preregistration into the theses writing



Source: Own compilation

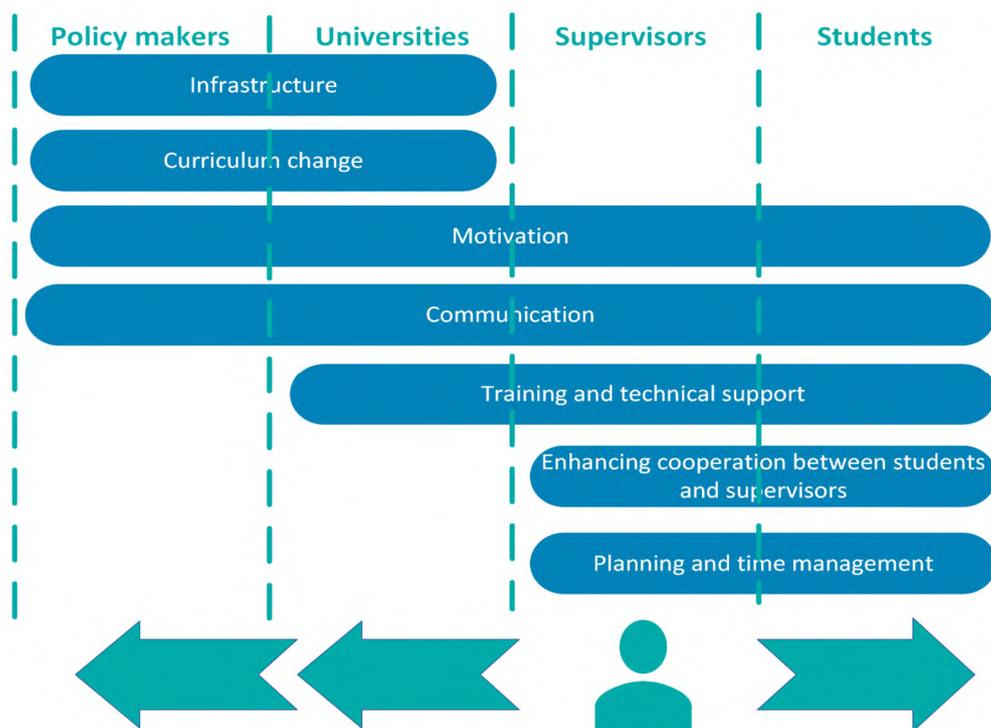
Preregistration could be complemented by a data management plan that students would create simultaneously as preregistration. This would teach students important data management practices in line with FAIR principles. At the same time, if students would collect unique data, it would be helpful to open it afterward so that other students or researchers could use it. However, data management is a very time-consuming activity, and I am not sure if the students could do it within the timeframe in which they are doing the final paper.

The exact placement in the theses writing timeframe will depend on the overall academic year setting at the institution. Button (2018), for example, offers a timeframe for the United Kingdom, where pre-registration work begins during the summer break, and then students give their first presentations of their pre-registration at the start of their final year of undergraduate study (October). Data collection then runs from November to March. In April, students discuss the main results of their studies and, after feedback, write the final version of their theses.

### The challenges of incorporating preregistration into theses writing process

Every change in the educational system is a complex and long-term process. For that reason, I present some of the challenges different stakeholders would face if preregistration became an official part of theses writing (see Figure 2). I present the challenges for three levels of stakeholders – policymakers, universities, supervisors, and students.

Figure 2 Key stakeholders and challenges in the process of incorporating preregistration into theses writing



Source: Own compilation

Firstly, it is necessary to provide infrastructure to support and encourage pre-registration. It is possible to use off-the-shelf solutions, e.g., the Open Science Framework - but it is up to policymakers and universities to provide the necessary resources for those interested in preregistration. Second, if preregistration is going to become an official part of university education systems, it is necessary to introduce this change in university curricula. Such a change is unlikely to occur at the national or even global level. Therefore, I assume it is more likely that individual universities will be pioneers and begin to change their curricula.

Two challenges cut across all key stakeholders – motivation and communication. Without proper motivation to shift towards the trends of transparent and open research at the level of theses writing, the proposed change will not be possible. Students, supervisors, universities, and policymakers must be motivated to change the perception of the thesis writing paradigm. But how to motivate such a body of stakeholders? A general shift towards open and transparent science could motivate policymakers and universities. There are calls such as the Paris Call on Research Assessment (French Open Science Committee, 2022), Plan S (cOAlition S and European Science Foundation, 2021), The Hong Kong Principles for assessing researchers: Fostering research integrity (Moher et al., 2020), Coalition for Advancing Research Assessment (COARA, 2022) or UNESCO recommendation on Open Science (UNESCO, 2021), which provide a direction for new practices in research assessment and day-to-day scientific work in line with principles for open and transparent research. The European Commission is a significant force for open and transparent research in Europe, requiring open research practices in funded projects (European Commission, 2019). In the U.S., a recent White House memorandum directs U.S. funders to modify their policies to ensure public access policies as soon as possible and no later than the beginning of 2026 (White House Office of Science and Technology Policy (OSTP), 2022). The National Institutes of Health is even faster in this regard and will require public access to research outputs for funded projects as early as 2023 (National Institutes of Health, 2022). These steps are slower or faster to trickle down to the national and local levels of individual states, research organizations, and universities. I believe that these activities will boost the principles of open science to the level of individual organizations and will thus gradually become a natural part of scientific practice (e.g., by the management of universities promoting the principles of openness, including pre-registration - both morally and at the level of science and research assessment, and the writing of theses).

Another important aspect is proper communication, which is critical at all levels of the proposed model. It is a crucial challenge to explain the importance of pre-registration research and how it can be done. Even though I mention communication as a challenge that cuts across all the levels of stakeholders mentioned, I believe that primarily librarians and theses supervisors will play a key role by communicating the importance of preregistration to university administration and guiding students to preregister their theses

Related to communication is the need to provide proper training and technical support. Again, here would, libraries have a significant role. Librarians will serve as the big wheel that will provide the necessary energy for the process I am proposing here. With the introduction of pre-registration for theses, there will be a great demand for the technical and a knowledge base that librarians offer. Libraries must have proper support from universities so that they can provide support to theses supervisors and students.

However, the most tremendous responsibility will remain with the supervisors and students. Great importance is the reinforcement of the time allocation for supervision of theses. Theses supervisors must have sufficient time to devote to the students. Preregistration is a time-consuming activity that requires a great deal of time commitment from both the supervisor and the student. Therefore, planning and time management are key factors in incorporating preregistration into the theses writing process. Also, as for introducing preregistration into theses writing, I believe this process will begin with individual theses supervisors leading their students through examples of good practice to preregister their research (as the position of persona in Figure 2 suggest). At the same time, as I mentioned above, I see a critical role for librarians who will become guides for researchers and students in the area of preregistration.

#### **Call for action**

Instead of a summary, I offer a call for action, which I believe will be more effective than a simple summary of what has been said above. Preregistration is a research roadmap, a

data-marked, predetermined, read-only research plan that, in most cases, is created before the study itself and published in a public repository. In its simplest form, pre-registration may involve only the registration of the basic study design, but it is possible to create a very detailed pre-specification of the study procedures, preliminary outcomes, and statistical analysis plan (Munafò *et al.*, 2017). A preregistration is a vital tool for promoting transparent, reproducible, and open research (Nosek *et al.*, 2015, 2018; Munafò *et al.*, 2017; Grahe, 2021). In the paper presented, I argue that for several reasons, preregistration should become a routine part of theses writing.

First, preregistration increased the credibility, transparency, and robustness of results. Pre-registration can alleviate frustration, both on the supervisor and the student, by clarifying what the research design is in advance. Preregistration also has educational benefits - getting students interested in data and data analysis early in the research process will increase their knowledge of data analysis and management and research procedures. Students will be thinking about their research design right at the beginning of the writing process, allowing them to avoid mistakes in their research approach. It will also make it 'easier' for students in the later stages of theses writing as all decisions about the research approach will have already been made. Students can also use preregistration as a reminder in cases where they forget a step of the proposed research design. This sort of reminder also reduces self-deception, which is another benefit of preregistration. Last but not least, preregistration is an excellent tool for communicating research and demonstrating a certain level of knowledge and skills that can increase the status of students in the research community.

In addition to the benefits of preregistration, it is essential to note that several major research funders, such as the European Commission, the White House Office of Science and Technology Policy, and the National Institutes of Health require the principles of open, transparent research. Thus, there is a presumption that this will be the future of the whole research community.

For this reason, I encourage all readers, including policymakers, university management or staff, but especially theses supervisors, librarians, and students, to start re-registering research produced as part of theses writing. I also offer for consideration the idea of supplementing preregistration with a data management plan that would further expand the data work acquired by students following FAIR principles.

There are many guidelines, outlined procedures, templates, and examples that can help students and theses supervisors move towards standardizing preregistration as part of theses writing. Some of the best guidelines are those from the Center for Open Science and the Open Science Framework (Center for Open Science, no date). In addition to step-by-step instructions on how to preregister, these tutorials include templates for different types of preregistration. Tutorials are not only in text form but also in the form of YouTube videos, so even those who prefer videos over text will enjoy them. If you would prefer some more traditional resources, I recommend the book *Journey into Open Science and Research Transparency in Psychology* by Jon Grahe (2021), which describes in a very nice way how to make research transparent, including pre-registration. Grahe (2021) encourages readers to activity at the end of each chapter with a variety of practical exercises and provides supplementary materials on the Open Science Framework. If you find the resources overwhelming, you can always contact your university librarian, who will be happy to advise you and direct you to valuable resources to help you navigate pre-registration issues.

The present proposal has several limitations, such as the fact that it is not elaborated in detail, is not set in a specific scientific context, is not discipline-specific, and does not take into account the heavy day-to-day agenda of thesis supervisors. The paper also does not include a comprehensive review of all projects and initiatives that incorporate elements to enhance the reproducibility of science in undergraduate programmes. Despite

limitation, this paper offers one of the first frameworks for incorporating preregistration into the thesis writing process.

### Acknowledgement

I would like to thank the Association of Libraries of Czech Universities, which supports librarians from all over the Czech Republic not only in the field of research but also in further education and future career development. Without the support of the Association of Libraries of Czech Universities, my participation in this conference would not have been possible. The results and knowledge included herein have been also obtained owing to support from the following institutional grant. Internal grant agency of the Faculty of Economics and Management, Czech University of Life Sciences Prague, grant no. 2022A0008.

### References

- Alger, B.E. (2020) *Defense of the scientific hypothesis: from reproducibility crisis to big data*. New York, NY: Oxford University Press.
- Button, K. (2018) 'Reboot undergraduate courses for reproducibility', *Nature*, 561(7723), pp. 287–287. Available at: <https://doi.org/10.1038/d41586-018-06692-8>.
- Center for Open Science (no date) *Registered Reports, Details and Workflow*. Available at: <https://www.cos.io/initiatives/registered-reports> (Accessed: 1 December 2022).
- cOAlition S and European Science Foundation (2021) *Plan S - Making Full & Immediate Open Access a Reality*. Available at: [https://www.coalition-s.org/wp-content/uploads/2021/03/Plan-S\\_profile\\_March2021.pdf](https://www.coalition-s.org/wp-content/uploads/2021/03/Plan-S_profile_March2021.pdf).
- COARA (2022) *Agreement on Reforming Research Assessment*. Available at: [https://coara.eu/app/uploads/2022/09/2022\\_07\\_19\\_rra\\_agreement\\_final.pdf](https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf) (Accessed: 1 December 2022).
- Cockburn, A., Gutwin, C. and Dix, A. (2018) 'HARK No More: On the Preregistration of CHI Experiments', in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18: CHI Conference on Human Factors in Computing Systems*, Montreal QC Canada: ACM, pp. 1–12. Available at: <https://doi.org/10.1145/3173574.3173715>.
- Crchartier, ~ (2018) *The Accelerated CREP, Psychological Science Accelerator*. Available at: <https://psysciacc.org/2018/04/08/the-accelerated-crep/> (Accessed: 1 December 2022).
- European Commission (2019) *Open Science*. Available at: [https://research-and-innovation.ec.europa.eu/document/download/77d34a83-7ed9-4041-9fb2-3be3094f2319\\_en?filename=ec\\_rtd\\_factsheet-open-science\\_2019.pdf](https://research-and-innovation.ec.europa.eu/document/download/77d34a83-7ed9-4041-9fb2-3be3094f2319_en?filename=ec_rtd_factsheet-open-science_2019.pdf) (Accessed: 1 December 2022).
- French Open Science Committe (2022) 'Paris Call – OSEC 2022'. Available at: <https://osec2022.eu/paris-call/> (Accessed: 1 December 2022).
- Grahe, J. (2021) *A Journey into Open Science and Research Transparency in Psychology: A Journey through National Parks*. 1st edn. New York: Routledge. Available at: <https://doi.org/10.4324/9781003033851>.
- Lakens, D. (2019) *The Value of Preregistration for Psychological Science: A Conceptual Analysis*. preprint. PsyArXiv. Available at: <https://doi.org/10.31234/osf.io/jbh4w>.
- Moher, D. et al. (2020) 'The Hong Kong Principles for assessing researchers: Fostering research integrity', *PLOS Biology*, 18(7), p. e3000737. Available at: <https://doi.org/10.1371/journal.pbio.3000737>.
- Munafò, M.R. et al. (2017) 'A manifesto for reproducible science', *Nature Human Behaviour*, 1(1), p. 0021. Available at: <https://doi.org/10.1038/s41562-016-0021>.
- National Institutes of Health (2022) *NOT-OD-22-213: Supplemental Information to the NIH Policy for Data Management and Sharing: Protecting Privacy When Sharing Human Research Participant Data*. Available at: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-22-213.html> (Accessed: 1 December 2022).
- Nosek, B.A. et al. (2015) 'Promoting an open research culture', *Science*, 348(6242), pp. 1422–1425. Available at: <https://doi.org/10.1126/science.aab2374>.
- Nosek, B.A. et al. (2018) 'The preregistration revolution', *Proceedings of the National Academy of Sciences*, 115(11), pp. 2600–2606. Available at: <https://doi.org/10.1073/pnas.1708274114>.
- Nosek, B.A. et al. (2019) 'Preregistration Is Hard, And Worthwhile', *Trends in Cognitive Sciences*, 23(10), pp. 815–818. Available at: <https://doi.org/10.1016/j.tics.2019.07.009>.
- Quintana, D.S. (2015) 'From pre-registration to publication: A non-technical primer for conducting a meta-analysis to synthesize correlational data', *Frontiers in Psychology*, 6(OCT), pp. 1–9. Available at: <https://doi.org/10.3389/fpsyg.2015.01549>.
- Reproducibility Education in an Undergraduate Capstone Course* (2021). (TIER Spring Symposium on Instruction in Reproducible Research). Available at: [https://www.youtube.com/watch?time\\_continue=1792&v=5ExuqTkNNuU&feature=emb\\_title](https://www.youtube.com/watch?time_continue=1792&v=5ExuqTkNNuU&feature=emb_title) (Accessed: 1 December 2022).
- UNESCO (2021) *UNESCO Recommendation on Open Science*. UNESCO. Available at: <https://doi.org/10.54677/MNMH8546>.
- White House Office of Science and Technology Policy (OSTP) (2022) *Desirable Characteristics of Data Repositories for Federally Funded Research*. Executive Office of the President of the United States. Available at: <https://doi.org/10.5479/10088/113528>

# Slovak Centre of Scientific and Technical Information **SCSTI**

Achieve  
your goals  
with us



## INFORMATION SUPPORT OF SLOVAK SCIENCE

### SCIENTIFIC LIBRARY AND INFORMATION SERVICES

- technology and selected areas of natural and economic sciences
- electronic information sources and remote access
- depository library of OECD, EBRD and WIPO

### SUPPORT IN MANAGEMENT AND EVALUATION OF SCIENCE

- Central Registry of Publication Activities
- Central Registry of Art Works and Performance
- Central Registry of Theses and Dissertations and Antiplagiarism system
- Central information portal for research, development and innovation - CIP RDI >>>
- Slovak Current Research Information System

### SUPPORT OF TECHNOLOGY TRANSFER

- Technology Transfer Centre at SCSTI
- PATLIB centre

### POPULARISATION OF SCIENCE AND TECHNOLOGY

- National Centre for Popularisation of Science and Technology in Society

### IMPLEMENTATION OF PROJECTS

- National Information System Promoting Research and Development in Slovakia – Access to electronic information resources - NISPEZ
- Infrastructure for Research and Development - the Data Centre for Research and Development - DC VaV
- National Infrastructure for Supporting Technology Transfer in Slovakia - NITT SK
- Fostering Continuous Research and Technology Application - FORT
- Boosting innovation through capacity building and networking of science centres in the SEE region - SEE Science

## Patent files: Case study of digitalization in The National and University Library of Slovenia

Mojca Trtnik and Veronika Potočnik, National and University Library, Slovenia

### Video Presentation

<https://av.tib.eu/media/59867>

### Abstract:

*The Ephemera and Grey Literature Collection of the National and University Library holds patent files acquired by Slovenia in accordance with the law on legal deposit. The patent files are interesting both from a thematic point of view, as they cover various fields (from printing and bookbinding to industries), and from a historical point of view, since they date back to the time of the Kingdom of Serbs, Croats and Slovenes and later the Kingdom of Yugoslavia (from 1922 to 1942). We have recognised the importance and value of this type of material and, in order to make it most widely accessible, we developed a plan for cataloguing and digitising the patent files in the Digital Library of Slovenia. Due to staff shortages and the large-scale project (11,000 items or 71,772 pages), we decided to include about half of all the material in the first phase of the project. After reviewing the existing cataloguing rules, we decided on a new way of processing this type of material: the collective entry.*

*The cataloguing of the patent files was a challenging task and involved collaboration with the National Bibliographic Centre. The staff of the Collection of Ephemera and Grey Literature prepared records in the Slovenian Cobiss mutual catalogue, thus enabling users to make the first contact with the material. Cataloguing was followed by the digitisation phase and inclusion of metadata in the Digital Library of Slovenia.*

### Introduction

In 1970, the National and University Library launched the *Collection of Special Literature*, due to the growing influx of "special materials". The collection was named the *Collection of Special Type of Literature* which in modern times means the collection of grey literature. The collection included all materials and documents of legal deposits collected in previous years. As an increasing number of ephemeras were added to the collection, it was renamed the Ephemera and Grey Literature Collection. The collection is unique even today because it marks a certain period of time. When we look back in the past, we can appreciate the value and uniqueness of a moment when the material of countries that were emerging and disappearing was acquired due to the legal deposit. National borders have changed, but material has remained. In the past, the question of whether ephemera should be kept, has often been discussed. At the National and University Library, we have followed the fundamental principle of legal deposit, which states that "all printed materials, regardless of their content, form or size, are equally important and must be preserved – they bear witness to the time in which they were created. Therefore, the value of the legal deposit collections is not assessed in terms of the value of individual works, but in terms of the complexity of a collection" (Kodrič-Dačić, 2002). In the slovenika (1) collection, ephemera occupies a special place. "The legal deposit rules flooded the public scientific libraries with material that, at first glance, had no scientific value. However, it has gained special importance in the light of the new role of regional libraries after the March Revolution. The legal deposit has become the core mechanism for the collection of material and the creation of the most comprehensive collections in the Crown land. It contributes to the culture of the country, its cultural history, literature and the creation of a regional bibliography. Therefore, ephemera should also be included in the collection with no exceptions, just like seemingly irrelevant diaries, since they have a special value in the light of potential studies." (Kodrič-Dačić 2001, 150)

One of unique collections from the time of the Kingdom of Serbs, Croats and Slovenes is also the Collection of Patent Documents, which was created in times that were important both from a historical and cultural-economic point of view.

Given the specificity of the area of the former Yugoslavia, we will present a brief historical outline of the time and the territory in which the patent files in question were created.

---

This is the material in the Slovenian language, about Slovenia and Slovenes, by Slovenian authors, Slovenian publishers, printed in Slovenia.

### **Patent files of the Kingdom of Serbs, Croats and Slovenes and the Kingdom of Yugoslavia**

The Kingdom of Serbs, Croats and Slovenes (abbreviated SCS) was a state formed on December 1, 1918 by the unification of the State of Slovenes, Croats and Serbs and the Kingdom of Serbia. On October 3, 1929, it was renamed the Kingdom of Yugoslavia. Alexander Karađorđević, heir to the Serbian throne, announced "the unification of Serbia with the lands of the independent State of Slovenes, Croats and Serbs into a single Kingdom of Serbs, Croats and Slovenes".

The State of Slovenes, Croats and Serbs, which was created at the end of the First World War and the collapse of the Austro-Hungarian monarchy, brought a short-lived independence to Slovenes in less than a month, but the country was not internationally recognised. Its position in the international community was hampered by the fact that it was founded on part of the territory that in the past belonged to the Austro-Hungarian monarchy, which was defeated in the First World War.

In accordance with the London Agreement concluded by the Entente Powers with Italy in the middle of the World War I, the SCS was threatened by the territorial claims of Italy. The 1915 Treaty of London gave Italy territorial rights to Primorska, Istria and parts of Dalmatia in exchange for entering the war on the side of the Entente Powers.

On July 20, 1917, representatives of Slovenes, Croats and Serbs from Austria-Hungary, assembled in the Yugoslav Committee, and representatives of the Kingdom of Serbia, signed a political declaration on the establishment of a common state. The Corfu Declaration states that Serbs, Croats and Slovenes are one nation with three names, which will unite to form a kingdom under the leadership of the Karadjordjević Dynasty from Serbia that was on the side of winners at the end of the WW1. Before the proclamation, representatives of the State of Slovenes, Croats and Serbs and of the Kingdom of Serbia agreed at the Geneva Conference in November 1918 to unite the two countries on a federal basis, but later developments led towards centralisation.



The coat of arms of the Kingdom of Serbs, Croats and Slovenes was based on the Serbian royal coat of arms with a double-headed eagle, added by Croatian chessboard with twenty fields and Slovenian half-moon, facing upwards, with a single five-pointed star above it. In 1921 the coat of arms was changed to some extent, with three hexagonal stars above the Slovenian half-moon (SI\_ZAL\_ŠKL/0278, Ivan and Franja Tavčar, *Visoko pri Poljanah*, vol. 2, a. e. 11).

On 5 January 1921, the Decree on the Protection of Industrial Property was published in the Official Gazette of the Land Government for Slovenia, the Kingdom of Serbs and Croats. The "Administration for the Protection of the Industrial Property" as an independent state office was established. It had headquarters in Belgrade, and was in charge of patents applications. Its jurisdiction covered the entire territory of the Kingdom of Serbs, Croats and Slovenes. It operated under the Ministry of Trade and Industry.

"The roots of the word 'patent' go back to the 13th century, when the term 'Letter patent' was used in England. It referred to a document - a public letter, delivered by a ruler. However, it is generally acknowledged that the system of patent law originated in the Republic of Venice. On 19 March 1474, the Republic of Venice adopted the Patent Law which was the turning point of patent law and laid the foundations for modern patent law" (Cokan, 2009).

In the SCS, the patent procedure was performed by the Administration for the Protection of Industrial Property; procedures followed the same methods that apply to patents today, such as Article 8 of the Official Journal (1921), which states that "patent right is acquired in the way specified herein, according to which its holder may, for a limited period, exclusively use and exploit the new invention in the field of handicraft or industrial production, place or sold objects produced by the invention into circulation. Article 11 of the Official Journal also states that only the inventor, his/her heir or an assignee is entitled to a patent. The patent was valid for 15 years (Article 18) and the invention was published in the Bulletin of the Board for the Protection of Industrial Property. Annex 1 provides an overview of the patent categories and their respective sections as published in the Official Journal. The same classification of subject fields was followed by the National and University Library in cataloguing patent files.

Like other materials, the National and University Library acquired the patent files as part of the legal deposit collected by the national libraries of the former Yugoslavia. Legal deposit can be considered as one of the mechanisms to promote the flow of and access to information and knowledge, and making them accessible in the common country. The notion of legal deposit appeared shortly after the invention of printing - initially, its function was to control printing. By the 17th century, it was known and legalised in most European countries.

.In the territory of today's Slovenia, the 1807 regulation was the first to introduce legal deposit of every printed publication for lyceum and university libraries. Initially, legal deposit was aimed at education, but as society changed, it was to regulate censorship, too.

In general, grey literature has a special place in the slovenika. In 1857, the Ministry of Worship of God and Learning also defined arrangement and storing of such material in libraries: the material was arranged chronologically, placed in boxes or folders and marked (classified) with place and subject headings. This method of processing the material was implemented in the National and University Library until the first cataloguing rules and international standards.

#### **Digitisation at the National at University Library**

Digitisation at the National and University Library started in 1996. Between 1996 and 2003, several digitisation projects were carried out that were launched on different websites. Based on the findings of modern library and information science, the Digital Library was founded in 2003. The dLib.si portal was developed and launched and made public in November 2005. Between 2006 and 2008, several large-scale digitisation projects were carried involving digitisation of older library material. Digitised material from other libraries was also added to the dLib.si portal. The EOD service - digitisation on demand - has also been fully implemented, allowing users to order digital copies of older books from NUK collections.

In 2007, the most important project to date, "The Digital Library of Slovenia - dLib.si was launched; it was financed by the Norwegian Financial Mechanisms. Over the years, the dLib.si website has been improved several times and supplemented with additional user options.

In 2009 and 2010, the *dLib.si Plus* project was carried out, funded by the Ministry of Higher Education, Science and Technology and the European Development Funds.

During the project, new functions of the dLib.si portal were developed, tailored to the needs of the partner organisations and different user profiles. Possibilities to search for material by using a Thematic Browser, collections and geolocation have also been improved. During this period, the national aggregator of e-contents in the field of culture was also set up; it enables heritage institutions in Slovenia to submit metadata about their digital resources to Europeana, the web portal of European digital libraries (dLib, 2022).

#### **Digitization and Cataloguing of Patent Files**

The National and University Library acquired the patent files by means of legal deposit, also of the Kingdom of the SCS. The figures show that this is an impressive collection, comprising 11.000 patent files. This significant number required a thorough consideration - how to arrange and present such a remarkable "finding" to the public. Patent files were neither catalogued nor documented during the past period, due to the shortage of cataloguers. However, our predecessors had organised and protected them in suitable folders kept in depots.

**Picture 1: Folders of patent files**

Recognising the historical value of the collection of patent files, we decided in 2020 to digitise the collection and make it available to a wider public as a digital collection.

This paper presents the cataloguing and digitisation of the patent files.

Any large collection - in our case of the patent files, first of all requires a good work planning. In 2020, there were only two librarians employed in the Grey Literature Collection (hereafter PKG). Due to the volume of the patent files, a broader plan was required involving other colleagues and the Library's experts covering different domains, such as processing, digitisation and promotion of services as well. From a storeroom located in another area of Ljubljana, the files were transferred to the PKG's working premises. During identification process, the files were sorted according to their thematic topics, and a list of all patent files was made. The data was taken from the Official Gazette of the Kingdom of Serbs, Croats and Slovenes and of the Provincial Government for Slovenia, in which the patent groups were divided by numbers and within subclasses. This also facilitated cataloguing process and allowed processing of larger volumes of material within a single category as seen in the table below (Table 1).

**Table 1: Examples of cataloguing and digitisation of patent files:**

Patent class / Title in a catalogue record	Catalogued and digitised items
Colorants, lacquers, repaints, adhesives	146
Bleaching, colouring, fabric printing, finish coating	82
Grinding and ironing	12
Small products and smoking devices	28
Elevators	24
Explosion and heat engines, gas compression engines, spring motors and weight engines	296
Explosive materials and manufacture of lighters	6
Electrotechnics	1347
Electrical engineering (electrical lighting)	2
Gas production and gas lighting	33
Iron manufacturing	97
Iron manufacturing (Division 1)	19
Photography	59
Photography and cinema	87
Musical instruments	1
Fuel	170

Shipbuilding and maritime navigation	51
Construction of railways, roads and bridges	141
Construction of railways, roads and bridges (street and road cleaning)	5
Hydraulic motors and wind turbines	57
Alkali and alkaline-earth salts industry, salina solutions	20
Grease industry, candles, soaps, mineral oils	238
Soda and other large chemical industries	150
Clothing industry	1
Clothing industry, except hats and footwear industry	90
Instruments	189
Paper products, paper processing and advertisements	144
Manufacture and treatment of leather	54
Hat-making and production of felt	19
Ice and frost making, heat transfer and food storage	93
Paper production	27
Manufacture of sheet metal, metal tubes, pipe fittings and wire, as well as rolling mills	8
Chemical processes and instruments	49
Chemical processes and instruments (liquid strainers and straining presses)	4
Chemical processes and apparatus, if not classified in special classes	628
Chemical processing of metals	24
Ceramic, stone and cement industries	311
Slaughterhouses	9
Locking products and cash registers	135
Bookbinding	42
Manufacture of wooden parts for carriages, velocipedes and motor vehicles	376
Writing and drawing materials	58
Mechanical treatment of metals (except metal-rolling mills in general)	115
Mills	67
Health care	234
Wood treatment and preservation	157
Footwear	24
Fireplaces and stoves	181
Heating (general)	292
Tools and appliances not specifically indicated	170
Cutting tools, swords and daggers	36
Steam boilers with appliances (excluding fireboxes)	244
Steam engines (with circular and oscillating concrete and rotating cylinders)	1
Steam engines (permanent, locomotives and ship engines)	102
Bakery	106
Beer, brandy, wine, vinegar, yeast	102
Blast-furnace	246
Knitting, lace making, curling, trimmings	31
Agriculture and forestry, horticulture and viticulture, dairy farming, zootechnics	369
Household pots and equipment	275
Purification and preparation of ores, minerals (sludge and excreta treatment)	90
Towing	40
Tow fibres	52
Sugar and starch production	55
Control equipment, registration tools and automated cash registers	40
Lighting installations (excluding installations for electric and gas lighting)	26
Preparations for bottling	137
Making artificial fertiliser	70
Blowers and air ventilators	25

Rescue and fire-fighting equipment	90
Hand and travel devices	12
Horn, ivory, rubber, gutta-percha and other plastics	98
Mining	11
Suction and other pumps for liquids	86
Signalling (excluding railway installations of this type)	41
Sports, games, folk festivities	89
Glass	72
Compression mills	34
Firearms, ammunition, defence	640
Machine parts	351
Drying and throwing machines	44
Bristles	14
Sewing and embroidery	24
Printing, typewriters, stampings	124
Knitting	60
Tobacco	78
Water, water supply and sewerage	28
Exterior building construction	148
Ropemaking craft	12
Aircraft	220
Rail traffic	288
Food	155
<b>TOTAL</b>	<b>11.339</b>

The patent files were held in carton folders, which did not fully shield the files from dust particles. However, most of the patent files were preserved to the extent that no restorations were necessary, so we could start processing - cataloguing and inventorying them. Due to the size of the patent files collection and the complexity of the subject matter (classification into sections), we have decided to catalogue the files as collective entries.

Cataloguing of each individual patent file would be time-consuming. According to experience, processing of collective entries is faster, and an unlimited number of items can be processed at the same time. Collective entries also enable more creativity - a title created by a cataloguer can also be changed, if a new title more accurately describes the collection and if better meets users' search requirements. Thus, when to decide for processing in **collective entries**.

Collective entries are appropriate for material that is not linked bibliographically but has a common element - it can be an individual, an institution, an event, a process or time period, a subject, a language, a genre, etc. The main types of material that are appropriate for collective entries of are grey prints and other non-literary material – which according to one or more of the above criteria, constitute a certain entity, either have a common socio-historical significance (for example, in our case, the patent files), and /or written in less commonly used languages, and/or scripts (e.g. Macedonian and Albanian primary school textbooks), and various monographs, whose content can be identified by one or a few subject headings (e.g. popular editions of Bible stories in foreign languages and/or scripts), and by type (children's picture books without text). Collective entries gather material that users usually never search for by title or even by author, but always by subject. A record for a large number of units connected in one or other way, provides a more extensive information than records for individual headings - the latter, because of their large numbers and the lack of links between them, actually lose their value. With collective entries we create our own collection and through a detailed processing of the content make the material retrievable.

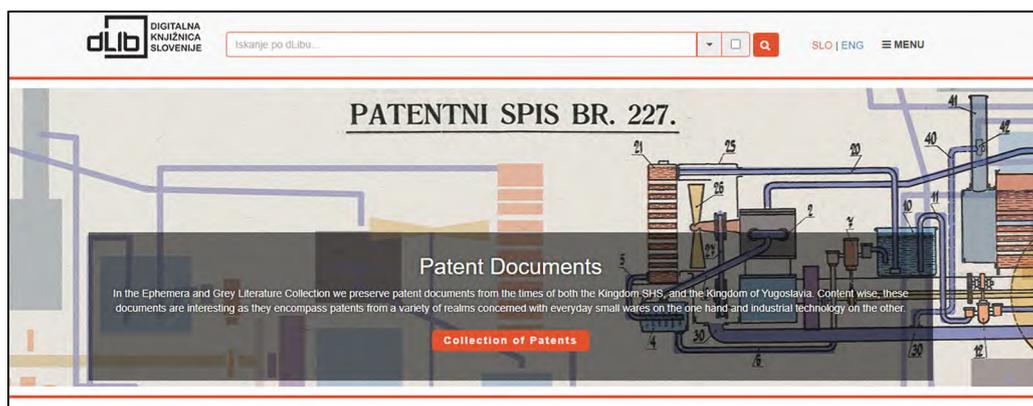
Table 2 provides a comparison and time assessment of two cataloguing techniques.

Table 2: Evaluation

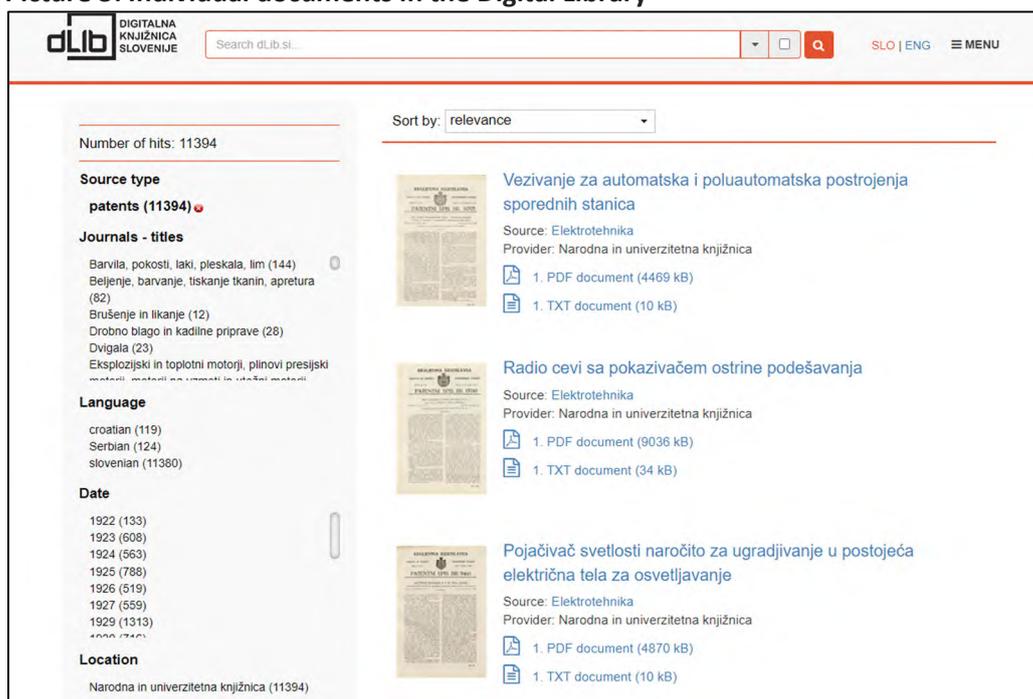
Type of record	Time of processing	Collecting period	Volume	Location	Standards	Title
Individual record	30 min	At once	1	Title, author	Rules	Static
Collection cataloguing	15 min	A period of time	Unlimited number of objects	content	Rules + creativity	Changing

During the project, a total of 61,772 pages of patent files were digitised and published on the Digital Library of Slovenia website.

Picture 2: Access to the digital collection of patent files



Picture 3: Individual documents in the Digital Library



Source: dLib, 2022

### Costs

The preparation of the material, bibliographic processing, organisation of digitisation, procurement of the procedures, submission and receipt of the material, publication were carried out by the staff of the National and University Library. The digitisation of patent files and scanogram processing were performed by an external contractor.

Estimated costs: EUR 25,000.

### Conclusion

Grey literature is published on an occasional basis and cannot be bought as part of the regular distribution network. To acquire as many of this material as possible, the National and University Library has decided to promote it. Grey literature is not traceable through regular book market; therefore, a correct promotion is an important step towards obtaining it. Of course, one might ask what is the purpose of collecting grey literature and ephemera. Nowadays, it is difficult to assess the value of a particular print as we don't know what the needs will be in the future. Its value increases with time, so it is important to capture the "moment" and create specific collections. Bibliographic processing of grey literature and ephemera requires a different approach and a wide knowledge of cataloguing. The best solution was to use collective entries, whose content definition is important, as it enables finding material quickly and easily.

We established that the decision on which material was suitable for processing with collective entries depended on a specific type of library and the size of its collections.

Digitalisation has completely changed the perception of grey literature and ephemera. That is the reason that the National and University Library has decided to digitise them and make as much material as possible available, while respecting copyright. This also makes the ephemera visible and valued differently.

The digitisation of patent files was in its final phase at the end of 2022 (the last 10,000 pages of patent files have been digitised). Most of the documents are currently available and accessible to the public on the Digital Library of Slovenia portal. In 2023, the project will be completed with the promotion and publication of the entire collection in the Digital Library of Slovenia.

### Bibliography

Cokan, P. (2009). *Od ideje do patenta*. [Magistrsko delo]. Nova Univerza, Evropska pravna fakulteta.

dLib. from <https://www.dlib.si/About.aspx>

Kodrič-Dačić, E. (2002). Slovenika. *Knjižnica*, 46(4), 65-85.

Potočnik, V., Velikonja, Š. (2017). Raznovrstnost drobnih tiskov : izkušnja v Narodni in univerzitetni knjižnici. *Šolska knjižnica*, 26(3/4), 5-17.

Pred sto leti je bila razglašena Kraljevina Srbov, Hrvatov in Slovencev (2018). from: <https://www.dnevnik.si/1042848618>

Uradni list Kraljevina Srbov, Hrvatov in Slovencev deželne vlade za Slovenijo (1921). Ljubljana, 5. januar, letnik III.

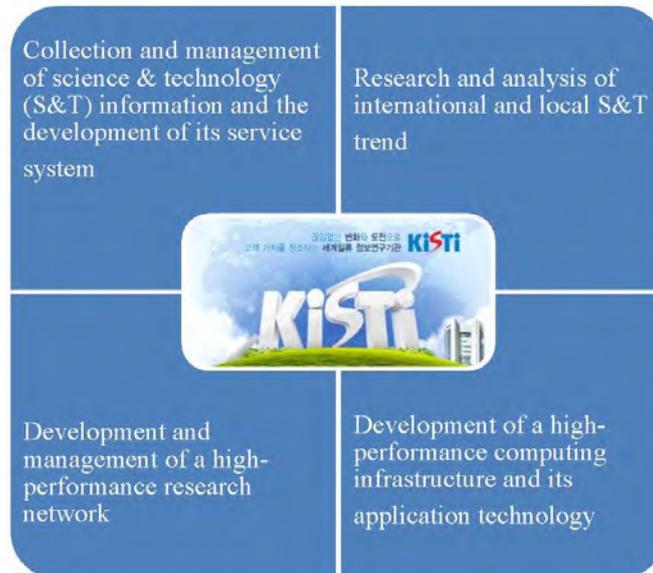
# Korea Institute of Science and Technology Information (KISTI)

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

## \* Vision

World-class information research institute creating values for customers

## \* Main functions



## \* Management and service of Korean R&D reports

KISTI exclusively manages, preserves, and serves Korean R&D reports for citizens and government officials. It provides Korean R&D reports and their information with National science & Technology Information Service (NTIS) and National Discovery for Science Leaders (NDSL).

## \* Contact information

KISTI email address: [hcpark@kisti.re.kr](mailto:hcpark@kisti.re.kr)

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

## Zine Making as Autoethnographic Serious Leisure

Andrea Marshall, Centre for Media and Celebrity Studies, United States

### Video Presentation

<https://av.tib.eu/media/59868>

### Introduction

Zines created in contemporary settings reflect the cultural practices that support a DIY (Do It Yourself) ethos and manifest as print and digital artifacts that resist construction as solely ephemeral. Hroch (2020) has observed that zines function as rebellious artifacts and that the act of making zines engage as processes that communicate radical ideas since “the post-digital situation makes us reconsider the categories of new and old media and radically remix them” (p. 22). Zines are hybrid media artifacts that demonstrate the attitudes and values of rich and varied communities of practice. The act of making zines is a counterhegemonic demonstration of resistance to orthodox principles of indexing information and collecting data since it is the pursuit of serious leisure.

Mansourian (2020) emphasizes serious leisure as a cultural practice that promotes the agency of amateurs and information seeking in serious leisure since “people involved in various forms of SL are usually passionate about what they do and it makes a big difference on the way they interact with information. They engage enthusiastically with the information they search for, find and disseminate.” (p.26). Furthermore, as “DIY is an amateur and participatory mode of engagement” (Watson and Bennett 2020 p.14) zines constantly disrupt traditional categories of data, information, and art through the deployment of bricolage, collage and montage as embodied practices that are adaptable and flexible. The open source tool the Electric Zine Maker promotes itself as a digital makerspace, as a learning environment and as a creative brainstorming tool. Digital zine making as DIY data creation and autoethnographic art curation asks questions about agency, identity and community. Zines as grey literature perform as beatnik boundary objects and remixed data artifacts that invite debate with regard to their content and their creators. Grey literature supports the unconventional with regard to knowledge infrastructures as well as information ecosystems. Zines challenge us to better understand how the DIY ethos can be applied to alternate ways of knowing to the access and the curation of information, and how to build data artifacts that can thrive in digital and analog environments.

### Zines as Boundary Objects and Data Artifacts

Zines as beatnik boundary objects dismantle extant power structures through surrealist critiques and bridge gaps between disparate ideologies through creating dynamic ways of understanding privilege and agency. J. Brett (2022) observes “zines are an ephemeral medium as much as any generalization can be made about zines and zine culture, they are creative endeavors of a particular time mindset and sociocultural milieu” (p.127). Zines as remixed data artifacts traverse the boundaries of digital and physical. In this way their ephemeral nature demands that we reimagine binary oppositions between the digital and the physical realms through the pursuit of serious leisure. Zines promote the practice of bricolage or learning by doing. Amateur zine makers collaborate with each other to learn about the world around them and build new understandings of the world around them. Through the application of bricolage collectively, these zine makers collectively construct new identities of expertise through the creation of zines as a communal practice.

J. Radway (2011) notes that “until zines emerged as digital forms they were generally defined as handmade non-commercial irregularly issued small run paper publications circulated by individuals participating in alternative special interest committees. Zines

exploded in popularity during the 1980s when punk music fans adopted the form as part of their do-it-yourself aesthetic and is an outside way to communicate among themselves about punks defiant response to the commercialism of mainstream society” (p. 148). Zines as punk rock boundary objects in montaged and collaged representations of ideologies mindsets and philosophies defy linear constructions of identity and categories of power. Zines as punk rock boundary objects resist one mode of understanding and one way of thinking. Rather, those that create zines invite those who read zines to become fellow collaborators and active participants in meaning making, while encouraging debates, discussions and differences. These interactive feedback processes evolve into varied values and perspectives within zine makers’ and zine readers’ communities of practice.

A. McNutt (2021) remarks that “zines are not a single monolithic genre. They are a medium which can contain countless narrative forms” (p. 1). This is because as S.M. Davidson (2022) ponders “ Zines thus cross a gap between life writing as a public act of visual narrative performativity via more formal modes of mass publication while nonetheless still engaging in dialogues concerning personal experiences with a more selectively created collective” (p.17).

Zines are by their nature informal and amateur, and their creation, consumption and publication all have roots in the practice of serious leisure. J. Radway (2011) comments on research about zines, how they function, how they are created and how they are consumed as questions that lead to more questions. She states “ indeed the preoccupation with what zines do and how they do it has become central to the burgeoning literature on zines a literature that has in effect created something that might usefully be termed “zine studies” and intellectual discourse about zines and zine ing that is not limited to the academic sphere” (J. Radway 2011 p.142). Zines bridge gaps by their very porous and informal montages of made meanings and collapse boundaries between privilege and resistance as well as formal knowledge structures and informal information ecosystems. Zines as serious leisure create informal expertise as a way to build communities. Zines as autoethnographic research create new information environments where the DIY ethos can be shared and become more innovative with more participation.

### **Zines as Grey Literature**

How do zines function as grey literature? M. Hroch (2020) states that “when speaking about the possibilities of printed matter zines are the most experimental platform because they are easy to create and cheap to print. In media theory zines are positioned in the alternative media paradigm” (p.21). Grey literature contains mutable bodies of content in which preprint publications, secondary drafts, white papers and multiple hybrid models of media exist. Zines as grey literature function as digital and analog artifacts that refute linear categorizations and question formal classifications. Zines are artifacts in motion; the meanings they make continuously shift with those who create them and those who consume them. M. Hroch (2020) furthers this notion “Zine communities are very material and intermaterial as various materials and printing machines are involved in the production and zine communities are therefore an assemblage of different bodies spaces objects machines and capital” (p. 21).

These “different bodies spaces objects machine and capital” (M. Hroch 2021 p.21) impact all of the identities and the ideologies of those who create them and those who consume them. The creation of zines and the consumption of zines are acts of defiance, acts of creativity and statements of identities. The feedback between those who create zines and those who consume zines is manifested in these artifacts. Zines are artifacts in motion and as the identities and ideologies that are associated with them are constantly evolving, zines are social movements.

### Zines as Social Movements

Zines are social movements that reflect the pursuit of serious leisure. A core ethos of zines as social movements is the DIY perspective. The do-it-yourself perspective of creativity and information seeking revels in informal learning, playful experimentation and the dismantling of power hierarchies that insist upon formalized modes of identity and information. Kuznetsov and Paulos (2010) define DIY perspectives as " we define DIY as any creation modification or repair of objects without the act of paid professionals. We use the term amateur not as a reflection on a hobbyist's skills which are often quite advanced but rather to emphasize that most of DIY culture is not motivated by commercial purposes" (p.1).

Watson and Bennett (2021) also observe that "one of the core aspects of zine culture is a DIY approach to practice" (p.121). The DIY ethos that informs the creation of zines and the construction of zine communities. This ethos has also infiltrated higher education research initiatives and research practices. Baker and Cantillon (2022) have remarked that " Zine making facilitates active participation and collaboration and offers a creative affective embodied way to contribute to research. Being amateur in style zines do not require any specialist skills there is no wrong way in terms of aesthetics to make a zine page with many forms of creative practice being appropriative but it is also reflexive and internal. Identities are remixed, repurposed and redefined within zine making. S.M. Davidson (2022) furthers this idea that "the zine artists' choice to not only refute the abstraction of limitations on creative production but also to operate themselves under a pretense of anonymity and of secrecy targeting the very technique of this expression of power. In doing so zinesters reflect the ability to create a self-sufficient literary economy existing in parallel to the culture capital machine" (p.17). Therefore, anonymity is agency in zine communities and zine identities are motivated by curiosity and playful experimentation. However this curiosity evolves with the zine creation and consumption into dynamic movements that create opportunities for reflexivity. This reflexivity generates dynamic bonds between members of zine communities and disrupt formal and social boundaries generating the potential for autoethnographic self-awareness.

Y. Mansourian (2022) furthers this notion by stating about serious leisure that " as you can see it is serious because they engaged with their chosen activity for the long term learn new skills and it can potentially form a new identity for them" (n.p.) Zine makers and readers are constantly investigating themselves, the world around them and their fellow non-conformist playful amateurs. Zine making as serious leisure, as autoethnographically reflexive participation and as learning is also a social movement. The DIY ethos is at the core of this social movement and S. Januario (2021) speculates about the DIY perspective as "once used as a means of indicating pockets of resistance to traditional forms of music and cultural production DIY has now become synonymous with a broader ethos of lifestyle policy that unites people and networks of alternative and trans local cultural production" (p.144). Zine making promotes the DIY ethos as a collaborative value that establishes recurring opportunities for information curation as creative bonding, autoethnographic reflexivity as evolving collective identities and remixed data artifacts as historical records of social change.

N. Stephens-Griffin and N. Griffin (2019) have observed that "autoethnography can be a means to amplify marginal or misrepresented voices or experiences" (p.3). Zine making as a social movement and as autoethnographic serious leisure continues to deconstruct barricades between both physical and digital environments. Zine making and zine reading support informal ways of knowing through creative collaboration and bricolage as a DIY practice.

### Zines as Open-Source Makerspaces

The open source tool the electric zine maker is at times an online makerspace and invitation to experiment with artistic impulses and an opportunity to experiment with bridging gaps between digital and physical environments. Open-source cultures support amateur interests that are collaborative and espouse innovative ways to adapt, remix and reuse digital tools to explore creative identities and meaning making. B. Ehn (2011) has described DIY pursuits and autoethnography through creative handcrafting when he states that “this kind of DIY auto-ethnography besides being enjoyable at least to myself also has the advantage of being doubly productive you manufactured lasting things with your hands as well as produce ideas” (p.59). Zine makers and consumers pursue creativity as a social movement, amateur identities as collaborative meaning-making as disruptive, defined and resistant to formal power hierarchies. Zine making promotes the exploration of values, identities and information that can bridge gaps and produce social change through creativity, curiosity and collaboration.

### Works Cited

- Baker, S., & Cantillon, Z. (2022). Zines as community archive. *Archival Science*, 1-23.
- Brett, J. (2022). " Zine Preservation", chapter in *Zines in Libraries: Selecting, Purchasing, and Processing*. American Library Association.
- Davidson, S. M. (2022). Pushing at the Seams: French Zines & Bricolage as a Liberatory Act.
- Ehn, B. (2011). Doing-it-yourself. *Ethnologia europaea*, 41(1).
- Januário, S.(2021). What if... the History told the resistance? The unlikely odds of the Do-It-Yourself attitude—insights regarding Zines and artistic heroes from the Oral Presentation Round Table 4 chair.
- Hroch, M. (2020). Not Out of Date, but Out of Time: The Materiality of Zines and Post-digital Memory. In *Forum Historiae. Časopis a portál pre históriu a pribuzné spoločenské vedy* (Vol. 14, No. 1, pp. 17-27). Historický ústav SAV.
- Kuznetsov, S., & Paulos, E. (2010, October). Rise of the expert amateur: DIY projects, communities, and cultures. In *Proceedings of the 6th Nordic conference on human-computer interaction: extending boundaries* (pp. 295-304).
- Mansourian, Y. (2020). How passionate people seek and share various forms of information in their serious leisure. *Journal of the Australian Library and Information Association*, 69(1), 17-30.
- Mansourian, Y. (2022). Information curation in serious leisure: How passion generates purposeful information activities.
- McNutt, A. (2021, October). On the Potential of Zines as a Medium for Visualization. In *2021 IEEE Visualization Conference (VIS)* (pp. 176-180). IEEE.
- <https://alienmelon.itch.io/electric-zine-maker>
- Radway, J. (2011). Zines, half-lives, and afterlives: On the temporalities of social and political change. *Pmla*, 126(1), 140-150.
- Stephens-Griffin, N., & Griffin, N. (2019, September). A millennial methodology? Autoethnographic research in do-it-yourself (DIY) punk and activist communities. In *Forum: qualitative social research* (Vol. 20, No. 3, p. 3). Institute for Qualitative Research.
- Watson, A., & Bennett, A. (2021). The felt value of reading zines. *American Journal of Cultural Sociology*, 9(2), 115-149.



the **leading** provider  
of out-of-the-box,  
**cloud repository** software  
for your **organization's**  
**grey literature**

visit [figshare.com](https://figshare.com)



- reports
- white papers
- government documents
- policy literature
- conference proceedings
- ...and more

10 years  
of figshare



## Digital Publishing, Open Access, and Grey Literature: The War in Ukraine 2022 as a Use Case

Dominic Farace, GreyNet International, Netherlands;  
Plato L. Smith, University of Florida; George A. Smathers Libraries, United States;  
Stefania Biagioni and Carlo Carlesi, InfraScience; ISTI-CNR, Italy

### Video Presentation

<https://av.tib.eu/media/59876>

### Introduction

The underlying approach in this study focuses on digital persistent identifiers and other linked open data as they become connected and interrelated in the course of research and whose outcome is published as grey literature. In January 2022, GreyNet published its 2<sup>nd</sup> edition of the International Directory of Organizations in Grey Literature (IDGL). This edition includes record entries from 45 countries worldwide with a listing of 224 organizations. Each entry contains the organization's URL and ROR ID. By way of the ROR ID, immediate access to other digital persistent identifiers is gained, such as GeoNames ID, CrossRef Funder ID, ISNI, and Wikidata. The digital publication of this directory provides a lead into our research project dealing with digital publishing, open access, and grey literature.

The War in Ukraine 2022 serves as GreyNet's use case. An online survey was designed and the questions formulated in such a way that a number of the responses provide additional linked open data and digital persistent identifiers derived from publications on the current War in Ukraine. These include URLs, DOIs, and ORCIDs. Survey data together with the linked open data gathered and compiled in this study are then analyzed. The results are expected to demonstrate the currentness of grey literature, its diverseness in formats and document types, the organizations that stand behind these publications, and how actionable persistent identifiers opens research in grey literature to a new level-playing field situated in a FAIR environment. An environment where data is not only findable and openly accessible but also interoperable and reusable by means of digital publishing.

### Method of Approach

A ten-question online survey was constructed and implemented via SurveyMonkey. Five of the questions were open-ended, three of which requested linked data. The other five questions allowed for Yes/No responses – four of which included comment fields. The first question on the survey required a Yes response. This confirmed that the survey respondent agreed to the Participant Consent Form stated at the start of the questionnaire that dealt with aspects of confidentiality, the risks and benefits, as well as the voluntary nature of the survey including a withdrawal clause.

### Survey Questions

1. Do you consent to take part in this study on the terms described above in the Participant Consent Form?
2. What is your organization's name and (if applicable) the acronym?
3. What is the URL of your organization's homepage?
4. Has your organization published one or more documents on the 2022 War in Ukraine?
5. If so, please provide a hyperlink or DOI to one of the publications?
6. What document type best describes the publication? (e.g., Blog, Podcast, Report, Policy Document, etc.)
7. Does the publication contain research data?

8. Does the publication in whole or part consist of audio-visual material?
9. Are you an author, co-author, or editor of the publication?
10. Please enter your name and/or ORCID along with your email address?

The link to the online survey was created on March 22<sup>nd</sup> 2022 less than a month from the start of the War in Ukraine. It was then posted to GreyNet's social media and distribution list on which all entries in IDGL are included. The online survey was closed on May 19<sup>th</sup> 2022 with 35 survey respondents. 88% of the survey questions were answered. And, on average, each of the 4 Yes/No questions were accompanied by 9.5 comments.

Responses to the ten survey questions were transferred to an Excel spreadsheet, whereby a record containing 19 fields was created for each of the 35 respondents. Six of these fields were added to each record derived from information and open data accessible to or assigned by GreyNet. These include the record ID, the organization's acronym, country of residence, sector of information, ROR ID, and whether it was included or not in the 2022 edition of the International Directory of Organizations in Grey Literature.

In brief, the respondents' answers to the online survey would allow for the capture of a number of types of linked data including: ORCIDs, DOIs, ROR IDs, URLs, as well as email addresses.

## Part One

### Overview of the Survey Findings

Of the 35 survey respondents four were anonymous. Fifteen of the respondents were from organizations in the academic sector, eleven from government, two from business, and the remaining seven were from NGO, IGO, or independent organizations. These organizations are located in fifteen countries worldwide: Australia, Bahamas, Canada, Czech Republic, France, Germany, Italy, Japan, Korea, Netherlands, Philippines, Slovakia, Slovenia, Spain, and the USA. Twenty-nine of these organizations carry a ROR ID and twenty-six were included in the International Directory of Organizations in Grey Literature (IDGL).

Twenty-three of the survey respondents indicated that their organizations had published about the War in Ukraine 2022, while eleven organizations had not yet published on the war as of the date on which they completed the online survey. One of the responses was considered invalid in that the linked provided was not related to the War in 2022.

The links to the publications were openly accessible and categorized in thirteen grey literature document types including: a survey <sup>(1)</sup>, website <sup>(4)</sup>, poster <sup>(1)</sup>, policy statement <sup>(4)</sup>, research guide <sup>(1)</sup>, panel discussion <sup>(1)</sup>, article <sup>(1)</sup>, blog <sup>(3)</sup>, report <sup>(2)</sup>, press release <sup>(1)</sup>, information note <sup>(1)</sup>, letter <sup>(1)</sup>, and speech <sup>(2)</sup>. Only three (13%) of the twenty-three publications contained research data, while eight (35%) contained audio-visual material. Four of the respondents were (co)authors or editors to one of the twenty-three publications. However, all twenty-three of the survey respondents have an assigned ORCID – indicating that they have publications on record.

### Summary of the Respondents' Comments

A summary of the 38 comments that accompanied the four yes/no questions on the survey is shown as follows:

When asked if their organization published one or more documents on the War in Ukraine 2022, those who answered 'yes', further commented that their library provided information services in Ukrainian as well as basic information about help for Ukrainian refugees in their country. Research grants were made available to Ukrainian researchers. Reports of statements by academic and religious spokesmen were published. Reliable curated links regarding the War in Ukraine were shared. Events were organized on

campuses dealing with the War. Appeals to end the War were drawn-up, signed, and published by both researchers and staff. Other respondents mentioned that while they did not publish, they had aggregated a manifold number of reports on the War in Ukraine.

Those who answered 'no' to the question responded that while their department had not yet published about the War in Ukraine, perhaps others on their campus did. Mention was made that technically they were not publishers. Further mention was made that their repository did however contain published articles dealing with that which led up to the War in Ukraine. Other respondents indicated that they intended to publish on the War in Ukraine in the near future.

When asked if the documents published on the War in Ukraine contained research data, three of the respondents confirmed that they did, while seven commented that their documents did not or that they were not aware of. One respondent mentioned that data was being compiled, but has not yet been published. Another commented that if reposts counted, then these did contain research data.

When asked if the publications contained audio-visual material, mention was made of video recordings of seminars and conferences dealing with the War in Ukraine. One respondent, who answered no to this question mentioned again that if reposts were included, then some of them contained audio-visual material.

In the final 'yes/no' question in the survey, when asked if the respondent was an author, co-author, or editor of the published documents, three stated that they were, another was a signatory to a publication, and again mention was made that his/her role was that of an intermediary and not an author.

### **Summary of the Primary Texts**

It is important to understand the chronological order in which the primary texts were published and later came to be included in this study. The key dates are the start of the War in Ukraine on the 24<sup>th</sup> of February 2022, the publication of the link to the online survey on the 22<sup>nd</sup> of March 2022, and the close of the link to the online survey on the 19<sup>th</sup> of May 2022. Hence, the survey was opened less than a month from the start of the war and closed just under three months into the War in Ukraine.

*An analysis of the twenty-three publications drawn from this study are grouped and summarized as follows:*

Clearly stated was a condemnation of the violence and an expression of shock. It was seen as a wake-up call, especially regarding the threat by Putin of nuclear deployment. This could lead to a sharing of nuclear weapons by Russia as well as Ukraine's allies.

Solidarity was called upon among Ukraine's allies and was witnessed and demonstrated in the way the citizens of Ukraine united in their response to the invasion. The principle of Sovereignty was upheld as contained in international law as well as in the U.N. Charter and reiterated by IFLA. By way of an EU Directive, the millions of Ukrainians that fled the war to neighboring EU countries were allowed to move through the EU as EU citizens. Likewise, Ukrainian research institutions were granted the same rights as those of the EU Member States. On the other hand, there were a suspension of agreements with Russia by European research centers.

The overwhelming and outspoken support for Ukrainian students can be understood given that near half of the survey respondents, who provided links to publications on the War in Ukraine were from academic institutions. This support is by way of scholarships, educational resources at no cost, and stipends. Support was also extended to staff and their families. Material support was further raised and mounted for refugees and other vulnerable groups affected by the War in Ukraine. This was voiced at campus gatherings.

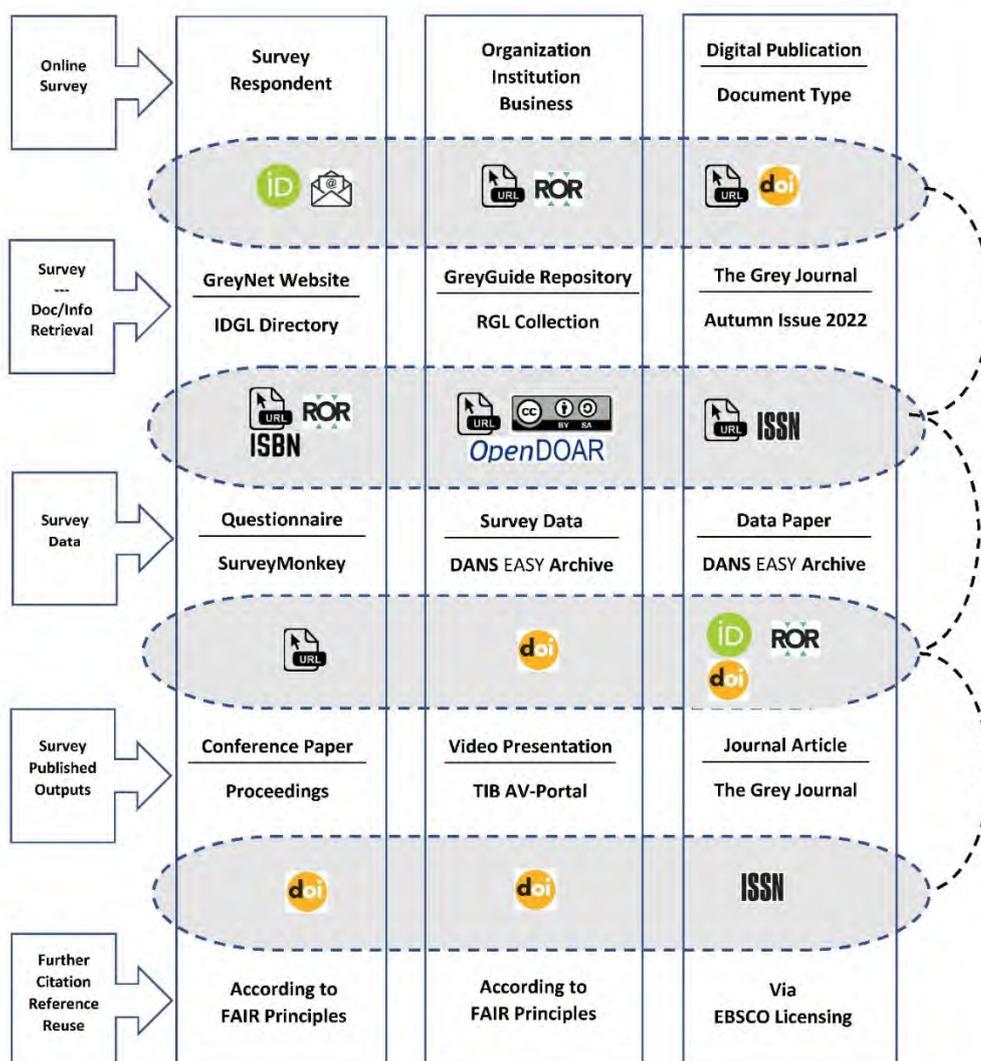
Express mention was made to avoid ad hoc policy decisions and to ensure accurate information about the War in Ukraine. While the free flow of information must be assured, it should not be at the cost of spreading disinformation and misinformation. A number of the publications addressed efforts to provide in-depth background information about that which has led up to the War in Ukraine. Mention was also made that print runs would be replaced by e-publications, due to the disruptions in deliveries and other information services.

Early on in the War, publications discussed the devastating effect it would have on the cultural heritage of Ukraine. Already a museum housing the work of an UNESCO recognized artist was destroyed. SUCHO, an initiative of over 1,500 international volunteers that collaborate online to digitize and preserve Ukrainian cultural heritage published a digital poster on their efforts. Likewise, early on in the conflict, FAO signals the risks involved for global agricultural markets that would aggravate the world food crisis.

**Part Two**

**Tracking Digital and Product Identifiers in Publishing**

The number of respondents to the online survey may not be considered a sizeable population in itself. However, when joined by persistent identifiers and other related data together with links to primary texts on the War in Ukraine, we can track and trace the impact this has for digital publishing and grey literature.



The schematic diagram (above) shows GreyNet’s publications both in textual and non-textual formats originating with the online survey and connected throughout by

actionable persistent identifiers and other product identifiers. Beginning with the publication of the link to the online survey (*top left column*), the responses provide access to data related to the respondents, their organizations, the types of grey literature documents published on the War in Ukraine, etc. In turn, these data are themselves connected to digital and product identifiers such as ORCID, ROR IDs, URLs, and DOIs.

### **Impact of Survey Data on Digital Publishing**

Data derived from the survey is then also available for inclusion in a revised and updated edition of the International Directory of Organizations in Grey Literature (IDGL). This information resource is openly accessible on GreyNet's website<sup>1</sup> and on the GreyGuide<sup>2</sup>, GreyNet's web access portal and repository listed in OpenDOAR. Further, by way of an advertorial in The Grey Journal<sup>3</sup>, greater awareness and wider coverage is given this grey literature resource. The URLs that link to IDGL and its assigned ISBN along with the ROR ID assigned to GreyNet<sup>4</sup> serve in implementing FAIR data principles<sup>5</sup>.

While the survey data remains accessible in SurveyMonkey<sup>6</sup>, it is further published in the DANS EASY Archive<sup>7</sup> and becomes openly accessible via CCO licensing. Added to the published survey data in DANS is a data paper<sup>8</sup> that provides a detailed and technical description of the data. This is further published in The Grey Journal licensed by EBSCO<sup>9</sup> and abstracted and indexed by Scopus and Clarivate. The published survey data and data paper carry with them persistent digital identifiers namely the DOI, ORCID, and ROR ID.

Now that the analysis of the survey data has been carried out, it awaits publication as a conference paper in the GLP Collection housed in the GreyGuide Repository<sup>10</sup>. It will be further compiled in the publication of the GL2022 Conference Proceedings<sup>11</sup>, and is further eligible for republication in a thematic issue of The Grey Journal. The video presentation of the conference paper will also be published in the TIB AV Portal<sup>12</sup>.

The survey data and accompanying data paper, the analysis of the survey data in the conference paper, and its video presentation each carry a separate but interoperable DOI. And, it is in this way that the survey data remain FAIR and open to citation, referencing, and reuse.

### **Concluding Observations**

It is important to note how prompt the organizations responded and with such foresight and concern for the accuracy in reporting on the War in Ukraine. The responses dealt with both the immediate consideration of housing for the millions of refugees fleeing Ukraine and for needed supplies, as well as considerations related to more lasting challenges in providing educational support for students and researchers and in dealing with the effects the war will have on world food shortages.

The survey provided an increase not only in GreyNet's digital content but also in new contacts. The fact that 74% of the respondents' organizations already appeared published in the International Directory of Organizations in Grey Literature (IDGL) confirmed that this resource was a determining factor in the survey population. The digital and product identifiers gained by the survey help in establishing the validity and reliability of the data collected. They demonstrate FAIR data principles in that they not only increase findability and ensure accessibility, but they also enable interoperability via the persistent identifiers and thus contribute in building the PID Graph<sup>13</sup> for further use in research and training.

Perhaps one of the most evident of observations is the presence of persistent and product identifiers in the trail of GreyNet's digital publications both in textual and non-textual formats. This provides an example of grey literature digital publishing, whereby the War in Ukraine served as a use case.

## Linked References

---

- <sup>1</sup> <https://greynet.org/internationaldirectory.html>
- <sup>2</sup> <https://doi.org/10.26069/greynet-2022-000.482-gg>
- <sup>3</sup> <https://www.greynet.org/thegreyjournal/currentissue.html>
- <sup>4</sup> <https://ror.org/01pxfxj80>
- <sup>5</sup> <https://www.go-fair.org/fair-principles/>
- <sup>6</sup> <https://surveymonkey.com/>
- <sup>7</sup> <https://doi.org/10.17026/dans-zhs-fewk>
- <sup>8</sup> <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:254710/tab/2>
- <sup>9</sup> <https://www.ebsco.com/products/research-databases/library-information-science-technology-abstracts-full-text>
- <sup>10</sup> <http://greyguiderep.isti.cnr.it/listtitoligl.php?authority=GreyGuide&collection=GLP&langver=en&RighePag=100>
- <sup>11</sup> <http://greyguide.isti.cnr.it/index.php/greyguideportal/document-share/gl-proceedings-1993>
- <sup>12</sup> <https://av.tib.eu/publisher/GreyNet%20International>
- <sup>13</sup> <https://www.sciencedirect.com/science/article/pii/S2666389920302440>

**ISTI (Institute of Information Science and Technologies "Alessandra Faedo")** is the largest institute for Computer Science of the **National Research Council of Italy (CNR)**. The Institute was founded in 2000 by merging the CNUCE and IEI institutes. The main aim of ISTI is to advance science and achieve important results in many Computer Science subdomains. ISTI is organized in the following thematic areas:

### **Networking Area**

This Area focuses on designing new algorithms and tools ranging from Smart Cities to intelligent Transportation Systems and from Ambient Assisted Living to monitor systems.

### **Software Area**

The Software Area performs third-party evaluation and certification of ICT processes and products, according to given requirements and standards and to meet the needs of users, industry, and public administrations. Industrial challenges, demands, and needs in systems and software development processes drive this thematic area.

### **Knowledge Area**

The mission of this Area is to investigate and advance the state-of-the-art in the Artificial Intelligence field, investigating applications to digital media and digital humanities and considering scalability issues. The Area addresses fundamental questions about the interaction between people and technologies and contributes to the evolution of science practices by researching, experimenting, and closely connecting research and development of innovative digital infrastructures, information systems, and intelligent (smart) solutions for fostering and empowering data centric research.

### **High Performance Computing Area**

The High-Performance Computing Area researches information indexing and retrieval, large-scale distributed/ cloud systems, big data analytics, machine learning and artificial intelligence, mobility analysis, and semantic enrichment.

### **Visual Area**

The goal of the Visual Area is to increase knowledge in signal processing, image understanding, and artificial vision in theoretical and applicative contexts. The goal is achieved by studying and developing model computer-aided methods, intelligent systems, machines for forming, elaborating, analyzing, and recognizing images and signals and their application in society, design of new computer graphics techniques, with a particular focus on geometry processing, learning, acquisition, visualization, and physical reproduction.

### **Flight and Structural Mechanics Area**

Flight and structural mechanics Area conducts research, software development, and consulting in continuum mechanics, mainly focusing on structural engineering. The main research topics are Computational solid mechanics, Mechanics of masonry structures and Structural health monitoring of historic buildings. The Area is also active in orbital debris modelling, mitigation and remediation, reentry predictions of uncontrolled spacecraft, and rocket bodies for civil protection applications, space experiments for fundamental physics, and mission analysis, including flight operational support and astrodynamics.

The ISTI Library was founded in 1954. Since 2000 is part of the Library and Scientific Documentation Center of the CNR Research Area in Pisa. The Library offers services supporting the scientific research of the CNR institutes located in the CNR Research Area of Pisa (approximately 2000 researchers and research support staff). The Library provides services to support the information needs of the scientific community. The Staff is responsible for the curation and enhancement of the institutes' heritage, facilitating access to either traditional or digital resources.

Among the Library's principal activities are managing the researcher's scientific production, supporting the principles of Open Science, and promoting free and open access to knowledge. The Library is involved with national and international associations and committees and collaborates with various partners in grey literature, digital libraries, institutional archives, open science, open access, and enhancement of bibliographic and documentary resources.

S. Giannini, S. Lombardi, A. Molino  
ISTI-CNR, Via G. Moruzzi, 1  
56125 Pisa, Italy  
Contacts: [library@area.pi.cnr.it](mailto:library@area.pi.cnr.it)



# Infrastructures for Science (InfraScience)

Connecting research and development  
to support Open Science

*Stefania Biagioni, Donatella Castelli*  
CNR-ISTI-Pisa Italy



## SCENARIO

Modern science is heavily data and compute-intensive, AI-assisted, participatory, and multidisciplinary. Similarly, sharing and publishing of scientific results are going to be revolutionized to support openness, transparency, and reproducibility, and to enable rewards for scientists who publish results of their work beyond the scientific articles. These approaches are expressions of a profound evolution of science practices that on the one hand is enabled by, and on the other demand for, continuous innovation in IT instruments and approaches.

## OUR MISSION

InfraScience is a Laboratory of the National Research Council of Italy - Institute of Information Science and Technologies (CNR - ISTI) based in Pisa, Italy. Its mission is to contribute to the evolution of data-centered research and open science practices and to foster and support the necessary corresponding change in the scientific communication approach.

## RESEARCH TOPICS

To support its mission the InfraScience Lab performs research in the areas of

*Data Infrastructures*

*eScience*

*Intelligent Systems*

Specific topics addressed include FAIR data management; Collaborative and social computing systems; Virtual Research Environments and Science Gateways; Scientometrics; Scholarly Knowledge Graphs; Computational reproducibility and provenance in scientific workflows; and Scientific publishing.

The Lab's mission is undertaken by investigating, experimenting, and closely connecting research and development in these areas to deliver innovative digital infrastructures and information systems. A major aim of the multi-areas integration is to foster and empower research workflows that are FAIR and open "by-design". By these properties, we mean first that processes and workflows are recognized as research products and second that, as such, they are automatically enriched with contextual metadata while they are executed. This approach lays the ground for a radical change in scientific communication. In the new vision, processes become first-class objects to be communicated and published. These objects offer a context that can be referred to by all the published results that are obtained in any of its steps. This relation largely contributes to explaining how they have been generated and thus it helps transparency and reproducibility.

## ACTIVITY & RESULTS

### FOUDED PROJECTS

[infrastructures.isti.cnr.it/projects](https://infrastructures.isti.cnr.it/projects)

### DATASETS

OpenAIRE Research Group  
OpenAIRE Covid-19

<https://graph.openaire.eu>

### INFRASTRUCTURE & SERVICES

D4Science, OpenAIRE,  
ISTI Open Portal,  
GreyGuide

[d4science.org](https://d4science.org)   [openaire.eu](https://openaire.eu)

[openportal.isti.cnr.it](https://openportal.isti.cnr.it)

[greyguide.isti.cnr.it](https://greyguide.isti.cnr.it)

### SOFTWARE

gCube, D-Net

[gcube-system.org](https://gcube-system.org)

[d-net.research-infrastructures.eu](https://d-net.research-infrastructures.eu)

### TRAINING ACTIVITIES & ORGANIZED EVENTS

WORKING GROUP  
TASK FORCES,  
&  
INTEREST GROUPS

[infrastructures.isti.cnr.it/](https://infrastructures.isti.cnr.it/) & [doi.org/10.32079/ISTI-AR-2022/001](https://doi.org/10.32079/ISTI-AR-2022/001)

## Grey is the new black: Changing library instruction virtually

Aleksandra Blake and Margaret McLeod, Carleton University, Canada

### Video Presentation

<https://av.tib.eu/media/59872>

#### **Abstract:**

*Searching for grey literature can often be a tricky and overwhelming process, in part because this topic is not always integrated into standard information literacy teaching sessions in post-secondary libraries (Mahood 2014, p.222). Despite the fact that grey literature is not considered scholarly, it is an important source of information for students, researchers and professionals in different areas of study and employment. This topic is often overlooked and not always integrated into standard information literacy teaching sessions in post-secondary libraries. And it should be. As Kingsley suggests, “the role libraries hold within research institutions is changing as the world shifts towards a digital and increasingly open future. This requires a rethink of the types of services and skill sets that are appropriate for an academic library to encompass” and teach (Kingsley, 2020, p.281).*

*While graduate students are well versed in searching for traditional academic literature (e.g., monographs and peer reviewed journal articles), they might be a bit “out of their league” in their ability to find grey literature (e.g., literature published outside of the mainstream academic and commercial publishing sectors). However, grey should be their new black— for example, many masters and doctoral students need to be equally capable of finding the material published by a wide range of researchers if they are to graduate with a robust set of professional skills that can be used in a variety of workplace contexts.*

*Our experience in designing and delivering a grey literature workshop tells us that our instincts are accurate—students are eager to learn these skills and put them to use in either a face-to-face or a virtual classroom.*

*Keywords: grey literature, library instruction, information literacy, graduate students, searching, databases, academic library, research, elearning*

#### **What is grey literature and why is it the new black?**

Grey literature (e.g., literature published outside of the mainstream academic and commercial publishing sectors) (Frater, Myohanen, Taylor & Keith 2007,146) is everywhere. Searching for grey literature (“grey lit”) is an important part of conducting systematic, scoping, or other comprehensive literature reviews in many disciplines. Access to grey literature allows researchers to find alternative perspectives that may not be represented in standard literature and locate experts in a particular field. In addition, grey literature offers the potential to balance any tendencies for bias found in mainstream published literature. Publications that can be defined as grey literature (for example: government documents, conference papers, clinical trials and technical reports, etc.) are timely, rapidly produced, and very current, offering an alternative to traditional academic peer reviewed sources.

The research suggests that citing grey literature has become more and more standardized. Scholars from different subject areas consider those resources valuable and helpful in bringing the academic community closer to civil engagement and matters important and relevant to societies (Kousha, Thelwall and Bickley, 2022). In addition, researchers need to find databases and repositories outside of traditional academic sources. Due to that fact researchers are encouraged to document and share their work and study strategies and processes involved in knowledge synthesis. Documenting the process can help others in discovery of grey literature by other researchers and contribute to greater exposure of

those resources to a wider community of scholars who are new to the field (Farrah & Mierzejewski - Urban, 2019). It is not a secret that at least one third of the reported research gets formally published, the other two thirds face delays in publications and become unavailable or discoverable via traditional academic channels (von Elm et al., 2033).

Kingsley suggests that evidence of impact of research outside of the academy “can take many forms, from annual reports from organizations and NGOs), minutes of community meetings...Given that librarians, by their profession, possess expertise in the location of information, they can offer support and advice to researchers need to collection this material” (Kingley, 2020, p.286)

Farrah and Mierzejewski - Urban (2019) argue that traditional databases often index resources that cover information about existing and outdated research. New and emerging trends are often missed and not considered or properly added to be discovered by researchers. In addition, the Kousha, Thelwall and Bickley (2022) research affirms that resources such as grey literature are becoming relevant in the academic environment and researchers should give those resources deserved attention. Traditional academic resources and non-traditional ones are becoming “equal” partners in the field of citations, “advantage of non- academic citations is that they can provide bridge to the external world and ground academic research in practical concerns” (Kousha, Thelwall and Bickley, 2022, p 3490).

In addition, this opinion is supported by current studies completed by Farrah and Mierzejewski- Urban (2019) which indicate that grey literature resources are playing a very important role in research that covers new and emerging health related technologies. Authors reviewed 22 reports published by 8 different organizations and institutes and concluded that almost half of the resources referenced in those documents were considered grey literature, which indicated that more and more scholars view these types of resources as relevant and valuable. Resources that came from the manufacturing industry, clinical trial registries, regulatory agencies and news agencies were only a few listed examples. This indicates the growing relevance and importance of resources that come from non-traditional publishing.

Students and faculty are often lost during the research in the vast world of grey lit. An important factor which hampers students from using and searching for grey literature is Information Fatigue Syndrome (IFS). IFS is defined as “A weariness or overwhelming feeling of being faced with an indigestible or incomprehensible amount of information (Savic, slide 4). This stress can “interfere with our sleep, sabotaging our concentration and undermining our immune systems” (Savic, slide 3). The sheer amount of grey literature available out in the universe is overwhelming and can be difficult to find. GreyNet International lists 150 types of documents that qualify as grey literature (2022). That is staggering. This is where librarians and subject specialists can play an important role and partly why we decided this was an important presentation for students. Using our grey literature Subject Guide, we can help by filtering information, listing high impact resources, and determining the reliability and trustworthiness of a source. Add to this the well-documented issue of stress on all university students, and you have a melting pot of problems that librarians can definitely help with.

Once students actually find grey literature, especially if it is Internet-based, there will be yet another factor that needs to come into play and perhaps add to the stress of academics and using grey literature. Although all resources used by students should be looked at through a critical lens, this is especially true of grey literature. “Information source assessment creates a new mental pathway that intervenes in the ladder of inference process and prevents us from forming misguided beliefs” (Lui, 76). And so begins the use of the CRAAP (Currency, Relevance, Authority, Accuracy, and Purpose) test for grey lit sources by evaluating any bias, relevancy, authority etc. There are other



Fortunately, we found that we agree on certain fundamental principles. In particular, we both believe that the primary job of an academic library instructor is to facilitate the discovery process rather than to provide answers to all the questions. As such, we think that active learning and access to specialized software are critical to creating a productive learning environment and teaching core skills such as strategic searching and evaluation of sources. The role of a research instructor is not only limited to pointing users to the right resources but also to teach them good research habits and skills that they will be able to apply in their future endeavors. Our instruction programs are designed to familiarize patrons with a physical and virtual collection and ways of retrieving both print and electronic resources. Moreover, the sessions we create teach students the jargon and vocabulary related to the field of subject related library research. We want to make sure that students know how to dissect the research process and to make it a more systematic process. Our goal is to ensure that the participants in our sessions gain knowledge that they will be able to take with them beyond their academic education. Plus, we want them to have fun!

In addition, we used the ACRL (Association of College & Research Libraries) Framework for information literacy in higher education as pillars in order to strengthen our pedagogical approach. Authority is Constructed and Contextual frame highlights to the researcher that it is an essential skill to “think critically about biases no matter the source” (ACRL, 2021). We believe that students need to be aware that think tanks, NGOs and any other interest groups produce relevant and important information that can be dependable and valid. Moreover, in our instruction we include ideas such as the Searching as Strategic Exploration frame “searching for information is often nonlinear and iterative, requiring the evaluation of a range of information sources and the mental flexibility to pursue alternate avenues as new understanding develops” (ACRL, 2021). As mentioned above, searching for grey literature requires creativity and critical thinking. Thus, we as instructors need to reinforce the idea that research for grey literature will challenge their perception of how and where to find information.

In February 2020, we offered this course face-to-face in the main library teaching space. Eleven grad students attended the session and we received positive feedback from the participants. For the purpose of this session, we used a previously existing grey lit help guide that contained essential information related to understanding and searching for these materials. We brainstormed ideas and came up with the following learning outcomes:

- Define what grey literature is and is not
- Justify why it is important to search grey literature when conducting knowledge syntheses
- Identify key grey literature sources
- Select grey literature sources specific to a topic and begin creating a search plan
- Articulate the process involved with producing a well-documented and transparent grey literature search

Engagement was high. We offered several activities such as small group work, grey literature recognition exercises, and open discussions. Positive responses from our participants encouraged us to continue our collaboration and to improve upon what we had already accomplished. The topics we covered included search strategies, identifying key authors, resource types and how to find them, web-searching tips which included Google Advanced tips and tricks, using search engines other than Google, using social media, finding grey literature in traditional library databases, keeping track of your searching and last but not least citation tips.

And then Covid hit. According to Miller and Janke (2022), “The most significant change during the pandemic was the rapid shift to online content and services...For many their instruction and consultations pre-pandemic had been a very in-person model.”

One of the respondents to their interview questions responded “noting the “biggest cultural shift” was to the completely virtual library” (p. 50). It was no different for us at the Carleton University library. It changed the way everyone needed to think, work and teach information literacy.

Moving forward required that we change our mode of delivery; the library building was closed as of mid-March 2020, which meant that we would have to switch to an online delivery format if we wanted to offer the workshop in either the summer or fall terms. Initially, we started this new design phase thinking that our online workshop would be less interactive for students. However, moving from in-person to online was easier than we thought. Although teaching via an online platform such as Zoom is different from being in the library classroom, we found that our students were engaged and just as interested as those who had taken the workshop in the early winter to the group of nineteen students in 2021 and twenty four students in 2022.

Our design process began with a review of the grey literature guide already available through the library website. We recognized that the current help guide was not the optimal teaching tool for an online workshop. Instead, we decided to create a subject guide that would be the basis for structuring our workshop and sharing information with students on an ongoing basis. Our subject guides are similar to Libguides but created on the Drupal platform.

We created a new subject guide for grey literature titled “Grey Literature” (<https://library.carleton.ca/guides/subject/grey-literature>, January 25, 2023) and taught the workshop using this guide because this type of online resource has greater visibility on the library website, and the layout of the guide allows for customization. We set up the guide to fit the desired Drupal-based format and organized the content based on a number of priorities:

1. The information needed to match the flow of our presentation.
2. The content needed to be readable online and easily discoverable by anyone who finds the guide without our support.
3. The information needed to meet the needs of our users, from novice to more experienced researchers. For example, the researcher is able to explore the different sections of the guide with no need to review the content if they are already familiar with specific sections.
4. The guide needed to reflect the steps that one needs to take while researching. First, they must familiarize themselves with the topic (what grey literature is and why it is relevant to explore it), then go on to the stage of searching (hence the guide, which directs users to links to databases and repositories), and finally they need to evaluate and cite information, which is the last part of the guide and our session.

Home / Guides / All Subject Guides /

## Grey Literature

Grey literature is defined as "information produced on all levels of government, academics, business and industry in electronic and print formats not controlled by commercial publishing" ie. where publishing is not the primary activity of the producing body."

ICGL Luxembourg definition, 1997. Expanded in New York, 2004. Grey literature can also mean literature that is hard to find or has inconsistent or missing bibliographic information.

- Getting Started +
- Search Strategies and Identifying Key Authors & Authorities +
- Search by Resource Type +
- Search the Web +
- Search in Library Databases +
- Keep Track of Your Searches +

**Contact**

**Margaret McLeod**  
  
 Subject Specialist  
 margaret.mcleod@carleton.ca  
 Schedule Appointment

**Aleksandra Blake**  
  
 Librarian  
 aleksandra.blake@carleton.ca  
 Schedule a Research Consultation

Figure 2: Webpage

This detailed subject (or "lib") guide replaces a three-page help guide. <https://library.carleton.ca/guides/subject/grey-literature>

We set up Zoom sessions for our presentations in the fall and winter terms 2021 and 2022. During the workshops, we presented the information by screen sharing the guide. Essentially, the guide was our instructional roadmap.

We had to work hard to keep up with the participants, because the Zoom chat box kept filling up with their answers and additional questions! We also set aside 15 minutes of the presentation time for more in-depth conversation by dividing participants into two breakout rooms and continuing to facilitate an open discussion. We had prepared some topics for analysis, but our participants had specific questions related to their own research and we followed their lead. It was interesting to watch how the discussion developed, especially the peer-to-peer collaboration. Thirty-one participants attended our one hour online session. In order to make the presentation more interactive, we used an application called Mentimeter, which is a web-based application that supports the creation of online polls with real-time feedback. In addition, we had the group do an interactive exercise in which they identified which items were grey or not grey. Padlet is another pedagogical tool that we used during presentations in order to interact with participants during the workshop. It is a cloud-based software, hosting a real-time collaborative web platform in which users can upload, organize, and share content to virtual bulletin boards called "padlets" (Appendix) . At the end our online courses were more interactive & multidimensional than our in-person session.

### **Debriefing: Grey is the new black and Zooming is the new interactivity**

We think that the online format made the session more interactive, and allowed the students to be more engaged with the multidimensional content than the in-person session. Participants seemed less nervous about contributing ideas online, especially through the chat feature. They were more likely to be on an equal footing than in the classroom, as everyone had a chance to respond whereas in class it might be only one or two who speak loudest and first. We applied active learning techniques by using

Mentimeter and Padlet to encourage engagement with the material and each other, screen sharing of documents, and online discussion rooms. All of these features allowed us to interact with students as much as possible and keep the conversation flowing.

Despite our initial apprehension, we are happy with the final product and students' reaction to our workshop. Moving from each section of the presentation was smooth and looked seamless, so this process was much easier for us than we had anticipated. Participant feedback was positive and overall, students enjoyed learning about Carleton's library technology to help with grey literature.

Most recent feedback included:

- "Really helpful to have the guide and to walk through the site in this way." (Anonymous, personal communication, November 2, 2022).
- "Very helpful, thank you! Love that the grey lit guide is available online so I don't have to remember everything or take notes." (Anonymous, personal communication, November 2, 2022).
- "Love the work the library has put into this. Thank you!" (Anonymous, personal communication, November 2, 2022).

Our next steps are to prepare to teach again in the winter term of 2023. We also hope to create an online module that students can explore on their own time to learn about grey literature. In addition, we will review the basic structure and content of the workshop and will update the subject guide to reflect comments and suggestions we have received from students as well as the constant changes in the field of grey literature.

In addition, based on our presentation we prepared a poster to be presented for the Twenty-Fourth International Conference on Grey Literature 2022. The poster takes you on a journey through the research process that leads to the discovery of grey literature and demonstrates how the library staff at Carleton University organized and delivered a graduate-level information literacy session entitled "Grey is the New Black: Searching for Grey Literature" in the new online reality. We shared our experiences of how we cultivated academic learning that is both interdisciplinary and overlooked in instructional practices.

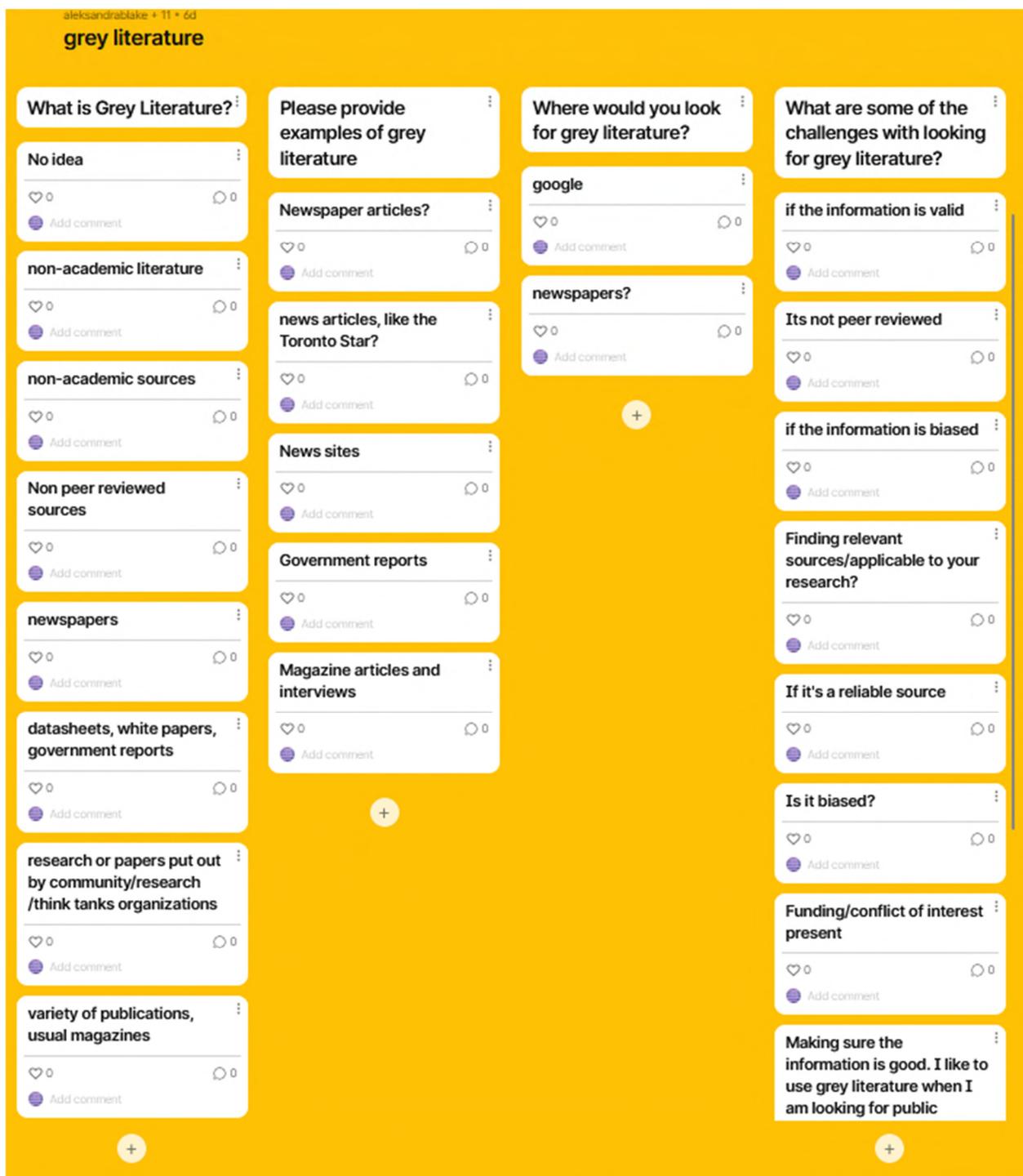
The importance of grey literature to both students and those in the workforce cannot be stated enough. The frustration and time commitment required to search for and evaluate these materials is critical. Engaging students in university with these materials is vital so that they have all the tools needed to satisfy both their academics and career goals. The library subject guides and workshops fulfill their need to learn how to find, evaluate and use these resources with the support of pedagogical resources such as the ACRL Frameworks. Research skills and emotional support should go hand in hand in order for young and up and coming researchers to be successful in their endeavors. In order to avoid Information Fatigue Syndrome (IFS) and feel confident in their own abilities in searching for grey literature successfully, the subject specialists need to continue to provide support.

Overall, we believe that by having to teach this workshop in person and online, we have strengthened our teaching praxis. In particular, by being able to provide information via different media (visual and auditory), we can reach a larger audience and enrich student learning. Plus, students continue to have the opportunity to learn a skill set that they can use professionally. We were able to support each other in this new quest by relying on our ongoing partnership and collaboration. In order to cultivate academic learning, it is vital to continually change our approach to teaching and implementation of our instructional practices. We were concerned about the switch to an online layout for the delivery of the workshop. In fact, moving from in-person to online was easier than we thought. We entered the project thinking it might be more difficult and less interactive, but found the students were engaged and just as interested.

## References

- ACRL, Association of College & Research Libraries. (2021). Companion document to the ACRL Framework for information literacy for higher education. Politics, policy, and international relations. [https://www.ala.org/acrl/sites/ala.org/acrl/files/content/standards/Framework\\_Companion\\_PPIR.pdf](https://www.ala.org/acrl/sites/ala.org/acrl/files/content/standards/Framework_Companion_PPIR.pdf)
- Bonato, S. (2018). *Searching the grey literature: A handbook for searching reports, working papers, and other unpublished research*. Rowman & Littlefield.
- Farrah, K. & Mierzejewski - Urban, M. (2019). Almost half of references in reports on new and emerging nondrug health technologies are grey literature. *Journal of Medical Library Association*, 107(1), 43-48. doi:dx.doi.org/10.5195/jmla.2019.539
- Frater, M. (2007). What would you tell me if I said grey literature? The New York Academy of Medicine grey literature report. *Journal of Electronic Resources in Medical Libraries*, 4(1-2), 145–153. [https://doi.org/10.1300/J383v04n01\\_13](https://doi.org/10.1300/J383v04n01_13)
- GreyNet International. (2022). *Document types in grey literature*. <https://www.greynet.org/greysourceindex/documenttypes.html>
- Kingsley, D. (2020). The 'impact opportunity' for academic libraries through grey literature. *The Serials Librarian*, 79(3-4), 281-289. <https://doi-org.proxy.library.carleton.ca/10.1080/0361526X.2020.1847744>
- Kousha, K., Thelwall, M. & Bickley, M. (2022). The high scholarly value of grey literature before and during Covid-19. *Scientometrics*, 127, 3489-3509. <https://doi.org/10.1007/s11192-022-04398-3>
- Liu, G. (2021). Moving up the ladder of source assessment. *College and Research Libraries News*, 82(2), 75-79. <https://doi:https://doi.org/10.5860/crln.82.2.75>
- Mahood, V. (2014). Searching for grey literature for systematic reviews: Challenges and benefits. *Research Synthesis Methods*, 5(3), 221–234. <https://doi.org/10.1002/jrsm.1106>
- Miller, K. & Janke, R. (2022). Canadian academic nursing librarians: Impacts of the COVID-19 pandemic on librarianship practice. *The Journal of the Canadian Health Libraries Association*, 43(2), 47-57. 10.29173/jchla29596
- Pappas, W. & Williams, I. (2011). Grey literature: Its emerging importance. *Journal of Hospital Librarianship*, 11(3), 228–234. <https://doi.org/10.1080/15323269.2011.587100>
- Savic, D. (2022, December 5). *Information Fatigue Syndrome and Digital Burnout*. [Conference presentation]. Twenty Fourth International Conference on Grey Literature, Amsterdam, Netherlands. <https://doi.org/10.5446/59863>
- von Elm, E., Costanza, M. C., Walder, B., & Tramèr, M. R. (2003). More insight into the fate of biomedical meeting abstracts: A systematic review. *BMC medical research methodology*, 3(12). <https://doi.org/10.1186/1471-2288-3-12>

Appendix



Actual responses from participants captured by the Padlet software.



# GreyGuide Portal and Repository

GreyNet's Flagship Project 2013-2022

Repository and Portal to Good Practices and Resources in Grey Literature

Stefania Biagioni and Carlo Carlesi, CNR-ISTI-Pisa, Italy  
 Dominic Farace, GreyNet International, Amsterdam, Netherland

Welcome to the GreyGuide, Your Open Access to Grey Literature Resources

<http://greyguide.isti.cnr.it>



GreyNet International together with the Institute of Information Science and Technologies of the National Research Council of Italy, **ISTI-CNR** mark this year a decade of collaboration in the design, development, and open access of diverse resources in the field of grey literature.

In this capacity, GreyNet acts as content provider and ISTI-CNR as developer and system provider for the GreyGuide Repository registered in OpenDOAR.

Together they have compiled and published the complete collection of conference papers in the GL Conference Series from 1993-2022. Further included are the accompanying abstracts, data papers, and biographical notes of hundreds of authors and researchers, who have contributed to this field of information.

The GreyGuide Portal further extends GreyNet's publication trail by providing open access to conference proceedings, slides, and posters via its **Document Share** section. The GreyGuide Portal further facilitates access to **Repositories** in which GreyNet's conference videos and published research datasets are housed.

## GreyNet's Publication Trail via the GreyGuide

GreyGuide Repository				GreyGuide Portal				
Proposals / Abstracts	Biographical Notes	Conference Papers	Data Papers	Conference Proceedings	Conference Slides	Conference Posters	Conference Videos	Published Datasets
GLA Collection	BIO Collection	GLP Collection	RGL Collection	Document Share	Document Share	Document Share	Repositories TIB AV Portal	Repositories DANS EASY Archive



In 2018, five years into this project, the GreyGuide came to include persistent digital identifiers (PIDS) namely the DOI, ORCID, and ROR ID in its metadata records. This has allowed for the crosslinking of the research data published in the DANS EASY Archive and the video presentations published in the TIB AV-Portal with their accompanying full-text conference papers published in the GreyGuide Repository.



DataCite Commons provides these actionable persistent identifiers with a single search interface and in so doing contributes to the construction of the PID Graph. The PID Graph makes it easier to describe complex use cases and relationships. It opens to new and further research via digital interoperable hyperlinks. And, it is in this way that the GreyGuide has enabled GreyNet to carry out research projects dealing with topics such as leveraging grey literature, policy development, data papers as trusted resources, securing open access, digital persistent identifiers, and digital publishing. In turn, this has qualified GreyNet as a research organization with a curated ROR ID.

### PARTNERS



<http://grey.net.org/>



CNR ISTI  
 InfraScience Lab  
<http://www.isti.cnr.it>

### CONTENT PROVIDER

Dominic Farace  
 GreyNet  
 International  
 Netherlands

### SYSTEM MANAGER

Stefania Biagioni  
 CNR-ISTI  
 InfraScience Lab  
 Italy

### TECHNICAL DEVELOPER

Carlo Carlesi  
 CNR-ISTI  
 InfraScience Lab  
 Italy

### COLLECTIONS

BIO: Who is in Grey Literature  
 GLA: Conference Abstracts and Proposals  
 GLP: Conference Papers  
 RGL: Resources in Grey Literature

### GreyGuide Rep

provides services to support the submission, description, searching, browsing, retrieval, access, preservation and visualization of multiple types of documents and multimedia.

Full compliant Open Access implementation via the OAI-PMH protocol → OpenAIRE



## GreyNet Membership 2023

### Choose your Level of Membership in GreyNet International

GreyNet International thanks its Associate, Institutional, and Individual Members for their continued support. It is through your membership that our host of information products, services, and resources can be sustained, enhanced, designed and developed. GreyNet International likewise takes this opportunity to invite other communities of practice dealing with grey literature to join us in 2023. Your membership is a guarantee to research, publication, open access, education, and public awareness of grey literature.

Associate Membership	Institutional Membership	Author • Individual Membership
<p><b>Associate Membership 2023</b></p> <p>Membership entitles Your Organization to Representation on the GL25 Program Committee</p> <p>Associate Membership further entitles your organization to the following conference benefits:</p> <p><b>GL25 Twenty-Fifth International Conference on Gray Literature 'Confronting Climate Change with Trusted Grey Resources'</b></p> <p>Member Information form</p>	<p><b>Institutional Membership 2023</b></p> <p>Institutional Membership includes information products, services, and advertising benefits</p> <p><b>GL25 Twenty-Fifth International Conference on Gray Literature 'Confronting Climate Change with Trusted Grey Resources'</b></p> <p>Member Information form</p>	<p><b>Individual Author Membership 2023</b></p> <p>Individual Author Membership includes information products, services, and considerable benefits</p> <p><b>GL25 Twenty-Fifth International Conference on Gray Literature 'Confronting Climate Change with Trusted Grey Resources'</b></p> <p>Contact Information form</p>

### An Overview of GreyNet's Information Products, Services, and Resources



## ***List of Participating Organizations***

Carleton University	Canada
Centre for Media and Celebrity Studies	USA
Consortium for Educational Communication	India
Czech University of Life Sciences Prague	Czech Republic
EBSCO Publishing	USA
Figshare	United Kingdom
Georgetown University	USA
GeoScienceWorld	USA
GERiiCO laboratory	France
German National Library of Science and Technology, TIB	Germany
GreyNet International	Netherlands
Information Today, Inc.	USA
InfraScience, Infrastructures for Science Laboratory	Italy
Institute of Information Science and Technologies, ISTI-CNR	Italy
International Council for Scientific and Technical Information, ICSTI	France
JPNA Trauma Centre, All India Institute of Medical Sciences	India
Korea Institute of Science and Technology Information, KISTI	South Korea
Leibniz Information Centre for Science and Technology	Germany
National and University Library, NUK	Slovenia
National Library of Medicine, NLM	USA
National Library of Technology, NTK	Czech Republic
Nuclear Information Section, International Atomic Energy Agency,	NIS-IAEA UN/Austria
Open Policy Ltd.	United Kingdom
Slovak Centre of Scientific and Technical Information	Slovakia
Smithsonian Libraries and Archives	USA
TextRelease	Netherlands
University of California	USA
University of Florida; George A. Smathers Libraries	USA
University of Lille	France
University of Wisconsin—Milwaukee	USA
U.S. Department of Energy; Office of Scientific and Technical Information	USA
Virginia Tech, University Libraries	USA
WorldWideScience Alliance	USA



# Twenty-Fifth International Conference on Grey Literature 'Confronting Climate Change with Trusted Grey Resources'

OBA Forum • November 13-14, 2023  
Oosterdokskade, Amsterdam, Netherlands

## Conference Announcement

With over a quarter century of research on grey literature carried out by diverse communities of practice in this field of information, a collective challenge emerges. Researchers and authors in sectors of government, non-government, academics, and business spanning manifold disciplines in science, technology, and the humanities are called to action. Their years of work dealing with the production, processing, digital publication, open access, and preservation of research outputs in multiple formats is called upon in confronting climate change.

At this point in time, with the advancements in information technology available to grey literature and in accordance with FAIR data principles, researchers, authors, librarians, and other information professionals and practitioners are tasked to ensure that research outputs are findable, accessible, interoperable, and render potential reuse in furthering research and education in their respective disciplines and sectors of information.

GL25 seeks to accept this challenge. To this end, grey literature communities worldwide are called upon to direct their attention in responding to climate change for the benefit of our vulnerable planet.

## Related Conference Topics

Climate Change • Sea Level and Coastal Environment • Fresh Water Resources • Fisheries and Aquaculture • Ecosystems • Biodiversity • Agriculture • Crops and Livestock • Food Production and Consumption • Weather Conditions • Natural Disasters • Public Health • Urbanization • Transportation • Greenhouse Gas Emissions • Energy Supply and Use • Solar Radiation • National Security • Social Policies • Climate Risks • Climate Strategies • Prioritizing Change • Climate Research • Sustainable Development • Capacity Building • Climate Science • Social Divide • Cultural Change • Vulnerable Populations • Ethical Issues • Climate Justice • Global Governance • Political Will • Etcetera

## Conference Dateline 2023

March 31 <sup>st</sup>	April 14 <sup>th</sup>	April 21 <sup>st</sup>	May 1 <sup>st</sup>	September 30 <sup>th</sup>	October 15 <sup>th</sup>	November 13 <sup>th</sup> - 14 <sup>th</sup>
Call for Papers Closes	Program Committee Meeting	Author and Session Notifications	Call for Posters Opens	Early Registration Closes	Conference Papers and Posters Due	GL25 Conference Convenes

International Conference Series on Grey Literature  1992-2023

**TextRelease**

GL25 Program and Conference Bureau

Javastraat 194-hs, 1095CP Amsterdam, NL  
<https://textrelease.com>  
[conference@textrelease.com](mailto:conference@textrelease.com)  
+31 (0) 20 331.2420



# Twenty-Fifth International Conference on Grey Literature 'Confronting Climate Change with Trusted Grey Resources'

OBA Forum • November 13-14, 2023  
Oosterdokskade, Amsterdam, Netherlands

## Call for Papers

<b>Title of Paper:</b>	<b>Conference Topic(s):</b>
<b>Author Name(s):</b>	<b>Phone:</b>
<b>Organization(s):</b>	<b>Email:</b>
<b>Postal Address:</b>	<b>URL:</b>
<b>Postal Code – City:</b>	<b>Country:</b>

## Guidelines for Abstracts

Participants who seek to present a paper dealing with grey literature are invited to submit an English language abstract between 300-400 words. The abstract should address the problem/goal, the research method/procedure, an indication of costs related to the project, as well as the anticipated results of the research. The abstract should likewise include the title of the proposed paper, conference topic(s) most suited to the paper, name(s) of the author(s), and full address information. Abstracts are the only tangible source that allows the Program Committee to guarantee the content and balance in the conference program. Every effort should be made to reflect the content of your work in the abstract submitted. Abstracts not in compliance with the guidelines will be returned to the author for revision.

## Related Conference Topics

Climate Change • Sea Level and Coastal Environment • Fresh Water Resources • Fisheries and Aquaculture • Ecosystems • Biodiversity • Agriculture • Crops and Livestock • Food Production and Consumption • Weather Conditions • Natural Disasters • Public Health • Urbanization • Transportation • Greenhouse Gas Emissions • Energy Supply and Use • Solar Radiation • National Security • Social Policies • Climate Risks • Climate Strategies • Prioritizing Change • Climate Research • Sustainable Development • Capacity Building • Climate Science • Social Divide • Cultural Change • Vulnerable Populations • Ethical Issues • Climate Justice • Global Governance • Political Will • Etcetera

## Due Date and Format for Submission

Abstracts in MS Word must be emailed to [conference@textrelease.com](mailto:conference@textrelease.com) on or before March 31<sup>st</sup> 2023. The author will receive verification upon its receipt. Shortly after the Program Committee meets in April, the authors will be notified of their place on the conference program. This notice will be accompanied by further guidelines for submission of full text papers, biographical notes, accompanying research data, PowerPoint slides, and required Author Registration.

## Author Information

**Bales, Brian**

11

Brian is currently the Coordinator of the International Nuclear Information System of the IAEA in Vienna. He worked previously for the International Residual Mechanism for Criminal Tribunals in Arusha, Tanzania, as well as a previous posting at the IAEA. He holds a Master's Degree in Information Technology from the University of Denver.

ORCID iD <https://orcid.org/0000-0003-1452-2785>

Email: [b.bales@iaea.org](mailto:b.bales@iaea.org)

**Biagioni, Stefania**

122

Stefania Biagioni graduated in Italian Language and Literature and specialized in Data Processing and DBMS at the University of Pisa. She is currently an associate member of the research staff at the Institute of Information Science and Technologies "Alessandro Faedo" (ISTI), an institute of the Italian National Research Council (CNR) located in Pisa. She is currently involved in the activities of the ISTI Infrastructures for Science Laboratory (InfraScience). She has been Head Librarian till August 2017. She was responsible of National and International Projects on Digital Libraries and Open Access Repositories looking at the DRIVER/OpenAIRE vision, such as ERCIM Technical Reference Digital Library (ETRD) and PUMA (PUBlication MAnagement) & MetaPub Projects. She has coauthored a number of publications dealing with digital libraries and grey literature. Her research interest is focused on digital libraries, knowledge sharing and transfer in scientific area, scholarly communication infrastructures, data and documents management, Open Access and Open Science. She has been dealing with grey literature since the 90's. In 2013, she became involved in the GreyGuide Project, where she now serves as its Repository and Web Portal Manager. In 2018 together with GreyNet International, DOIs were minted and entered in the metadata fields of GreyGuide records. Stefania's involvement with persistent identifiers carried over to the AccessGrey Project in 2019. She is currently involved in the PID Project that includes the ORCID and ROR ID together with DOI persistent identifiers in metadata records housed in the GreyGuide Repository.

ROR\_ID <https://ror.org/05kacka20>

ORCID iD <https://orcid.org/0000-0001-9518-0267>

Email: [stefania.biagioni@isti.cnr.it](mailto:stefania.biagioni@isti.cnr.it)

**Blake, Aleksandra**

130

Aleksandra Blake is a subject specialist in Research Support Services at the Carleton University Library, providing research support in areas of Political Science, International Affairs, Political Economy and European and Russian Studies. She holds a Master of Library and Information Science degree from San Jose State University, and a Bachelor's degree in Political Science from Carleton University. Her research interests include information literacy, accessibility of government information, and qualitative research.

**Carlesi, Carlo**

122

Carlo Carlesi graduated in Computer Science and has worked since 1970 at the IEI (now ISTI) of the Italian National Research Council in Pisa. His interest has been concentrated in many software engineering fields such as: Data Base Systems, Software quality and testing, Administration and UNIX system management, Network security and management, Data Publication, Open Science, Digital Library [Management] Systems and Architectures. He was Head of the "Information Technology Security Service" of the Institute and has participated in many national and international projects,

among the most recent: Italian Project for Research in the Antarctica - South Pole Project with the aim to create a Multidisciplinary Integrated Information System. Other of the projects in which he participated include: ERCIM Technical Reference Digital Library - ETRDL Project; PUMA (PUBlication MAnagement) a software infrastructure for the management of the institutional/disciplinary archives of the published and grey literature of the CNR Institutes. He is currently a collaborator of the Infrastructures for Science Laboratory (InfraScience) of the CNR ISTI and is involved in the GreyGuide Repository and Portal to Good Practices and Resources in Grey Literature. He has co-authored a number of publications dealing with digital libraries and grey literature. In 2013, he became involved in the GreyGuide Project, where he now provides for the Technical Development and Support. In 2018 together with GreyNet International, DOIs were minted and entered in the metadata fields of GreyGuide records. Carlo's involvement with persistent identifiers carried over to the AccessGrey Project in 2019. He is currently involved in the PID Project that includes the ORCID and ROR ID together with DOI persistent identifiers in metadata records housed in the GreyGuide Repository.

ROR\_ID <https://ror.org/05kacka20>

ORCID iD <https://orcid.org/0000-0001-9808-6268>

Email: [carlo.carlesi@isti.cnr.it](mailto:carlo.carlesi@isti.cnr.it)

**Farace, Dominic**

122

Dominic Farace is Head of GreyNet International and Director of TextRelease, an independent information bureau specializing in grey literature and networked information. He holds degrees in sociology from Creighton University (BA) and the University of New Orleans (MA). His doctoral dissertation in social sciences is from the University of Utrecht, The Netherlands, where he has lived and worked since 1976. After six years heading the Department of Documentary Information at the Royal Netherlands Academy of Arts and Sciences (SWIDOC/KNAW), Farace founded GreyNet, Grey Literature Network Service in 1992. He has since been responsible for the International Conference Series on Grey Literature (1993-2013). In this capacity, he also serves as Program and Conference Director as well as managing editor of the Conference Proceedings. He is editor of The Grey Journal and provides workshops and training in the field of grey literature.

ROR\_ID <https://ror.org/01pxfxj80>

ORCID\_ID <https://orcid.org/0000-0003-2561-3631>

Email: [info@grey.net](mailto:info@grey.net)

**Marshall, Andrea**

117

Andrea Marshall is a librarian and feminist researcher who has published work in human computer interaction, information science, feminist approaches to maker culture and STEM education and fan studies. Her current work includes gendered analyses of online Star Wars fanfiction communities and zines as grey literature.

ORCID\_ID <http://orcid.org/0000-0003-1229-4694>

Email: [andrearieethmarshall@gmail.com](mailto:andrearieethmarshall@gmail.com)

## Author Information *CONTINUED*

### McLeod, Margaret

130

Margaret McLeod is a subject specialist in Research Support Services at the Carleton University Library, providing research support in areas of African studies, social work, and global and international studies. She is a Carleton History graduate with interest in information literacy, primary source research, and library teaching/instruction.

### Plank, Margret

83

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.

ROR\_ID <https://ror.org/04aj4c181>

ORCID\_ID <https://orcid.org/0000-0001-8941-7563>

Email: [margret.plank@tib.eu](mailto:margret.plank@tib.eu)

### Potočnik, Veronika

107

Veronika Potočnik: Studies of Librarianship and Information Science; graduation univ. dipl. bibl. 2000; Specialist exam for Librarian Specialist (2009); Study of basic conduct of mediation process (2013); Employment at Faculty of electrical science - library (1999-2004); Employment at National and University Library (2005-); Head of Collection of ephemeral print and grey literature (2010-2020); Cooperation in Association of librarians in Ljubljana as the member of the board (2011-2015); Regular lectures to the students in Librarianship and Information Science in University of Ljubljana about cataloguing the ephemeral print (2012-); Cooperation in the organization committee for library conferences ZBDS (2005 and 2007); Mentoring to the students in Librarianship and Information Science in University of Ljubljana (2010-2014). Main fields of research (keywords): ephemeral prints, grey literature, librarianship, cultural heritage.

ROR\_ID (URL): <https://ror.org/04tjx1w16>

Email: [veronika.potocnik@nuk.uni-lj.si](mailto:veronika.potocnik@nuk.uni-lj.si)

### Prost, H el ene

33

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

ROR\_ID <https://ror.org/02mn0vt57>

ORCID\_ID <https://orcid.org/0000-0002-7982-2765>

Email: [helene.prost@inist.fr](mailto:helene.prost@inist.fr)

### Savi , Dobrica

25

Dr. Dobrica Savi  is information and knowledge management consultant and the former head, the Nuclear Information Section, 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 with management and library, information, and knowledge management services across various United Nations agencies, including UNV, UNESCO, the World Bank, ICAO, and the IAEA. His main interests are digital transformation, grey literature, creativity, and innovation.

ORCID\_ID <https://orcid.org/0000-0003-1123-9693>

Contact: [www.linkedin.com/in/dobricasavic](http://www.linkedin.com/in/dobricasavic)

### Sch opfel, Joachim

33

Joachim Sch opfel is associate professor in information and communication sciences at the University of Lille, member of the GERiiCO laboratory and independent consultant at the Ourouk consulting office, Paris. He is interested in scientific information, academic publishing, open science and grey literature. He is a member of GreyNet, euroCRIS and ND LTD.

ROR\_ID <https://ror.org/02kzqn938>

ORCID\_ID <https://orcid.org/0000-0002-4000-807X>

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

### Smith, Plato L.

49, 122

Plato Smith is the Data Management Librarian at the University of Florida with experience in academic research libraries, digital libraries, and data management. He received his doctorate in the field of Information Science from the School of Information within the College of Communication and Information at Florida State University, Florida's iSchool, Summer 2014. From 2005 to 2012, he was Department Head for the FSU Libraries' Digital Library where he developed, populated, and managed digital collections in the FSU Libraries' digital content management system, DigiNole Repository, and electronic theses and dissertations (ETDs) institutional repository.

ROR\_ID <https://ror.org/02y3ad647>

ORCID\_ID <https://orcid.org/0000-0003-1814-0151>

# PUBGREY

Publishers in Grey Literature

Registry of organizations serving as publishers in the field grey literature

**PUBGREY** is a unique information resource connecting the supply and demand sides of grey literature

**PUBGREY** manifests that producers of grey literature are also its users

**PUBGREY** is as diverse in its coverage of document types as it is in the organizations that publish them

**PUBGREY** offers your organization the recognition it deserves as a publishing body

GreyNet International seeks to maintain a registry of organizations that serve as publishers in the field of grey literature. The PUBGREY Registry contains standardized records submitted by the publishing bodies. Records include the sector of information, document types, publication formats, main areas of coverage, submission and acceptance, open access compliance, contact details, and identifying characteristics of the organization. PUBGREY Registry is a unique information resource connecting both the supply and demand sides of grey literature. <https://www.surveymonkey.com/r/J8DJSF5>

## Sample Records

**cadth** Canadian Agency for Drugs and Technologies in Health

**Sector of Information:** Independent, not-for-profit organization  
**Document Types:** Reports, Health Technology Assessments  
**Publication Formats:** Online, PDF  
**Main Coverage:** Health Care, Health Care Systems, Drugs, Pharmaceuticals, Medical Devices, Medical Diagnostics, Medical Procedures  
**Content Contributors:** Internal only  
**Open Access Compliant:** Yes  
**Homepage URL:** <https://cadth.ca/>  
**Identifiers:** Orste founded 1383; <https://ror.org/02fw4cc98>; GRID grid #13289.5; <http://lanl.org/isni/0000000085833941>; <http://www.wikidata.org/wiki/Q5029662>; [https://twitter.com/CADTH\\_ACMTS](https://twitter.com/CADTH_ACMTS); <https://www.linkedin.com/company/cadth>  
**Point of Contact:** <https://www.cadth.ca/contact-us>

**UF** Digital Collections  
George A. Smathers Libraries; University of Florida

**Sector of Information:** Academics  
**Document Types:** EIDs, grant proposals, datasets, reports, maps  
**Publication Formats:** PDFs, statfiles, images  
**Main Coverage:** Agriculture, Entomology, Caribbean studies, Children's literature, Book arts, Florida history  
**Content Contributors:** Internal only (University affiliates)  
**Open Access Compliant:** Yes  
**Homepage URL:** <https://ufdc.ufl.edu/>  
**Identifiers:** Permanent URLs  
**Point of Contact:** Chelsea Johnston  
<https://orcid.org/0000-0001-9215-6843>  
[cjohnston@ufl.edu](mailto:cjohnston@ufl.edu)

**NDLTD** Networked Digital Library of Theses and Dissertations (NDLTD)

**Sector of Information:** Academics  
**Document Types:** Theses and Dissertations  
**Publication Formats:** PDF plus metadata plus other formats as provided by authors  
**Main Coverage:** All disciplines; topics, subjects covered in theses or dissertations  
**Content Contributors:** We harvest metadata from all interested providers  
**Open Access Compliant:** Free, open access  
**Homepage URL:** <http://ndltd.org>  
**Identifiers:** Incorporated 2003  
NDLTD Awards: <https://ndltd.org/etd-journal/>  
<http://theses.org>  
<http://dissertations.org>  
Social Media  
**Point of Contact:** Edward A. Fox  
<https://orcid.org/0000-0003-1447-6870>  
[fox@ndltd.org](mailto:fox@ndltd.org)

**IAEA** International Nuclear Information System (INIS)

**Sector of Information:** International Organization  
**Document Types:** Reports, Proceedings, Theses, Patents, Legislative Material, Standards, Videos  
**Publication Formats:** PDFs, Metadata, Online  
**Main Coverage:** All areas of nuclear science and technology, including nuclear energy, nuclear medicine, radiation physics, soil remediation, etc.  
**Content Contributors:** Internal and External  
**Open Access Compliant:** Yes  
**Homepage URL:** <https://inis.iaea.org/search>  
**Identifiers:** Founded 1970  
<https://www.iaea.org/resources/databases/inis>  
**Point of Contact:** Brian Bales  
<https://orcid.org/0000-0003-1452-2785>  
[b.bales@iaea.org](mailto:b.bales@iaea.org)

GreyNet

## *Index to Authors*

### ***A-B***

Adie, Euan	77
Bales, Brian	11
Barnes, Christopher	49
Biagioni, Stefania	122
Blake, Aleksandra	130
Bucknell, Terrence	77

### ***C - F***

Carlesi, Carlo	122
Deumens, Erik	49
Farace, Dominic	122

### ***G - J***

Glover, Jennifer	77
Jaso-Tamame, Ángel Luis	77

### ***L-M***

Lipinski, Tomas A.	61
Marshall, Andrea	117
McLeod, Margaret	130

### ***P***

Plank, Margret	83
Potočnik, Veronika	107
Prost, Hélène	33

### ***R-S-T***

Reece, Alistair	17
Savić, Dobrica	25
Schein, Laura Christine	61
Schöpfel, Joachim	33
Símová, Tereza	99
Smith, Plato L.	49, 122
Stooksbury, Shelby	83
Trtnik, Mojca	107



## An International Journal on Grey Literature

The Grey Journal is a flagship journal for the 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 also is itself a document type that is 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, such as researchers, editors, librarians, documentalists, archivists, journalists, intermediaries, etc.

### Subscription Order Form

For *GreyNet* Members free or reduced rates, see <https://greynet.org/membership.html>

<b>TGJ Volume 19, 2023</b>	Type of Subscription:	Amount in Euros	Total
THE GREY JOURNAL - PDF/ Email / PWP ISSN 1574-180X Annual Subscription, including Electronic Handling	<input type="checkbox"/> Institutional	€ 240	€

#### Customer information

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

#### Check Method of Payment:

Direct transfer to TextRelease, Rabobank Amsterdam, Netherlands  
BIC: RABONL2U IBAN: NL70 RABO 0313 5853 42, with reference to: TGJ/The Grey Journal

MasterCard/Eurocard       Visa card       American Express

Card No. \_\_\_\_\_ Expiration Date: \_\_\_\_\_

Print the name that appears on the credit card, here \_\_\_\_\_

Signature: \_\_\_\_\_ CVC II code: \_\_\_\_\_ (Last 3 digits on signature side of card)

Place: \_\_\_\_\_ Date: \_\_\_\_\_

**NOTE** Creditcard transactions authorized by Saferpay, Secure Paygate for VisaCard and MasterCard

#### Correspondence Address:

**TextRelease** Javastraat 194-HS, 1095 CP Amsterdam, Netherlands  
+31-(0) 20-331.2420 • [info@textrelease.com](mailto:info@textrelease.com) • <https://textrelease.com>