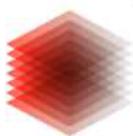


Conference Proceedings

ISSN 1386-2316



Program Sponsors and GreyNet Associate Members



TIB LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY

Nuclear Information Section

NIS - IAEA

International Atomic Energy Agency



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

NTK

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



EBSCO
PUBLISHING

UF George A. Smathers
Libraries
UNIVERSITY of FLORIDA

NLM
National Library of Medicine
United States



GL2021 Program and Conference Bureau

TextRelease

Javastraat 194-HS, 1095 CP Amsterdam, Netherlands
www.textrelease.com • conference@textrelease.com
Tel. +31 (0)20-331.2420



CIP

GL2021 Proceedings

Twenty-Third International Conference on Grey Literature “Digital Transformation of Grey Literature: Exploring Next Generation Grey”. - December 6, 2021 / compiled by D. Farace and J. Frantzen ; GreyNet International, Grey Literature Network Service. – Amsterdam : TextRelease, February 2022. – 132 p. – Author Index. – (GL Conference Series, ISSN 1386-2316 ; No. 23).

TIB (DE), CVTISR (SK), EBSCO (USA), ISTI-CNR (IT), KISTI (KR), NIS-IAEA (UN), NTK (CZ), NLM (USA), and the University of Florida (USA) are Corporate Authors and Associate Members of GreyNet International. These proceedings contain 12 full text conference papers presented during the plenary, panel, and poster sessions. Included is an author index with the names of contributing authors and researchers along with their biographical notes., ORCID, and ROR ID affiliation. A list of participating organizations as well as sponsored advertisements are likewise included.



Foreword

DIGITAL TRANSFORMATION OF GREY LITERATURE: EXPLORING NEXT GENERATION GREY

Grey literature is once again at a crossroads in its known 80 years in the vernacular. It appears that every quarter or more century this field of information faces a change in direction.

In the mid-20th Century, it originated with the collection of government documents – namely war reports. In the years succeeding, it expanded to numerous other types of documents and collections produced by organizations in government as well as academics and business. This period was occupied with the acquisition and document delivery of these scientific and technical materials. It was also during this period that the many problems associated with these materials surfaced, namely their lack of indexing, translation, open access, preservation, and assessed value for science and society.

Since the final decade of the 20th Century up to the present, the field of grey literature has initiated and undertaken evidenced based research hand in hand with technological developments and sustained information management. The problems of the prior period became challenges and this led to programs of research and education in grey literature. This period coincides with the digital transformation of grey literature. While much has been accomplished in connecting the supply and demand sides of grey literature, the field once again finds itself in transition driven by the fact that the entire information landscape is itself in a period of unprecedented change and flux.

GL2021 offers the many and diverse communities of practice in grey literature a unique opportunity to collaborate in addressing and defining the next phase in the digital transformation of grey literature. Together this can be accomplished by unlocking the potential next generation grey holds for information science and society.

Dominic Farace
GREYNET INTERNATIONAL

Amsterdam, Netherlands
FEBRUARY 2022

GL2021 Program Committee



Plato L. Smith ^{Chair}
University of Florida;
George A. Smathers
Libraries,
United States



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



Dobrica Savić
Nuclear Information Section,
International Atomic Energy
Agency, United Nations



Margret Plank
German National Library
of Science and Technology,
Germany



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



GL2021 Program Committee *CONTINUED*



Hana Vyčítalová
National Library of
Technology
Czech Republic



Joachim Schöpfel
University of Lille
France



Marta Dušková
Slovak Centre of Scientific
and Technical Information
Slovakia



David Baxter
University of Alberta
Canada



Stefania Biagioni
Infrastructures for Science
Laboratory, InfraScience
ISTI-CNR Italy



Dominic Farace
GreyNet International
Netherlands





GL2 021 Program and Conference Sponsors



TIB, Germany

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



NLM, USA

National Library of Medicine



ISTI, Italy

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



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



Table of Contents

	Foreword.....	3
	Program Committee	4
	Program and Conference Sponsors.....	6
	Conference Program.....	9
Program	I. Digital Transformation and the Changing Information Landscape.....	11
	II. Insights and Issues that Challenge and Demonstrate the Role of Grey Literature.....	55
	III. Exploring Next Generation Grey – Panel Survey.....	117
Advertorials	EBSCO Library, Information Science & Technology Abstracts with Full Text (LISTA).....	8
	INIS, The International Nuclear Information System.....	10
	CVTISR, Slovak Centre of Scientific and Technical Information.....	16
	TIB, German National Library of Science and Technology, Germany.....	32
	ISTI-CNR, Institute of Information Science and Technologies.....	54
	PsycEXTRA via EBSCO.....	66
	KISTI, Korea Institute of Science and Technology Information.....	84
	NTK, National Library of Technology, Czech Republic.....	110
	GreyGuide – An Example of Open Access publishing in Grey Literature.....	116
	GL2022 Conference Announcement.....	130
	Managing Grey Literature : Technical Services Perspectives.....	132
Appendices	List of Participating Organizations.....	125
	Author information.....	126
	Index to Authors.....	131

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

www.ebscohost.com



Conference Program

I. Digital Transformation and the Changing Information Landscape

- The Impact of Digital Transformation on the Sustainability of Grey Literature** 11
Dobrica Savić, Nuclear Information Section, International Atomic Energy Agency, NIS-IAEA, United Nations
- Grey Literature in Open Repositories: New Insights and New Issues** 17
Joachim Schöpfel and Eric Kergosien, University of Lille- GERiICO, France
Hélène Prost, CNRS - GERiICO, France; Florence Thiault, University of Rennes 2, France
- Improving guidelines for video abstracts: An analysis of the most popular videos abstracts in the TIB AV Portal** 33
Margret Plank and Jens Kösters, Leibniz Information Centre for Science and Technology, Germany
- Grey Literature and Persistent Identifiers: GreyNet's Use Case** 45
Dominic Farace, GreyNet International, Netherlands,
Stefania Biagioni and Carlo Carlesi, GreyGuide ISTI-CNR, Italy
Chris Baars, DANS-KNAW, Netherlands

II. Insights and Issues that Challenge and Demonstrate the Role of Grey Literature

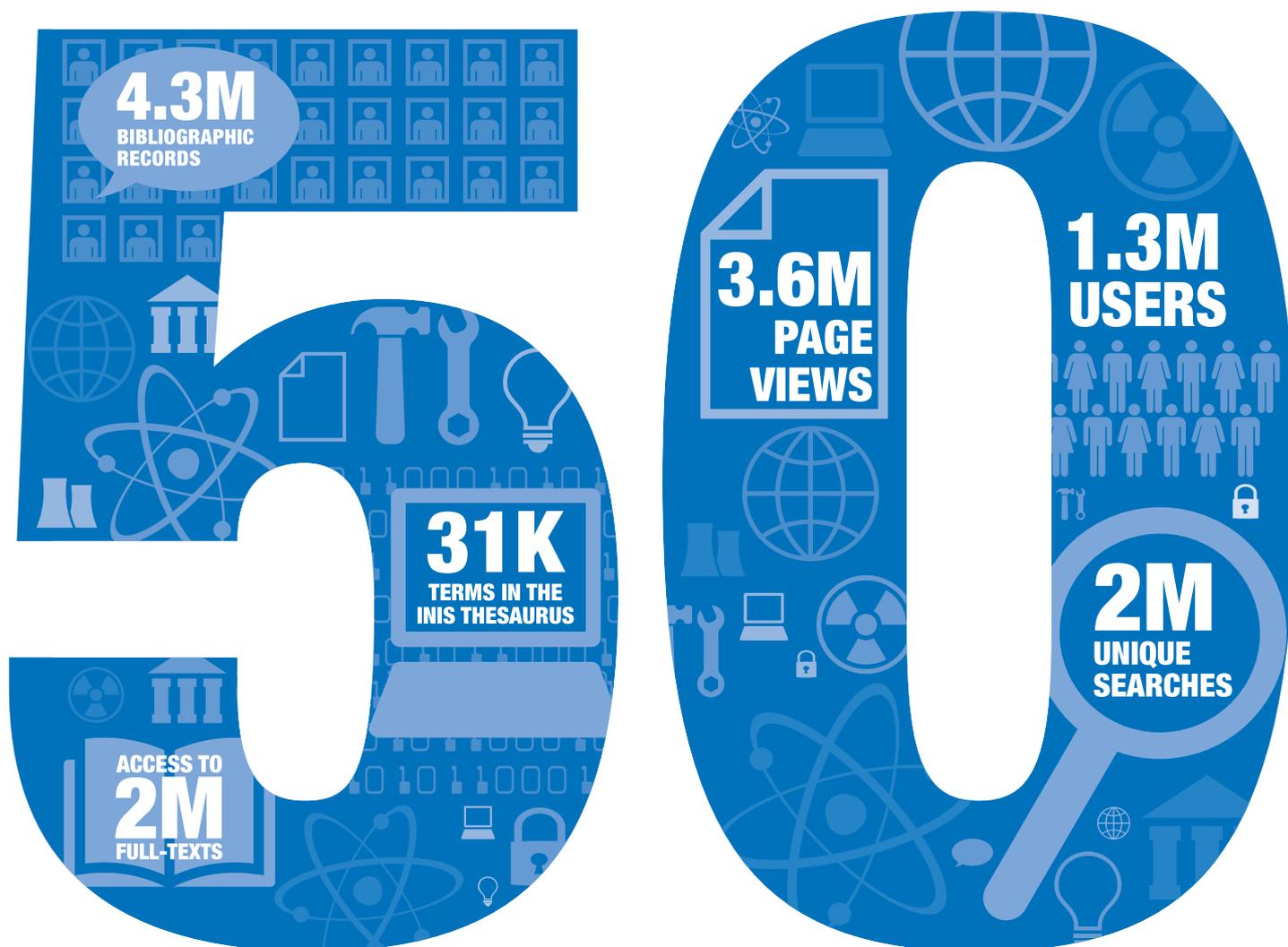
- Deepfakes: A Digital Transformation Leads to Misinformation** 55
Nika Nour and Julia Gelfand, University of California, Irvine United States
- Grey Literature citation and inclusion rates in gambling review articles: Opportunities for improvement** 67
David Baxter, Department of Political Science, University of Alberta, Canada
- Digital Grey Soviet Science** 85
Yuliya B. Balashova, Saint Petersburg State University, Russia
- Zines as Nonbinary Objects and Questions of Privilege** 89
Andrea Marshall, Centre for Media and Celebrity Studies, United States
- The endless life of OA Journals from myth to reality: Survey on the present status of vanished OA journals in Iran and future prospect** 93
Hamideh Memari Hanji, National Library and Archives of Iran
- Burning Grey: The Worldwide Influence of a Locally Published Grey Literature** 101
Vince Ervin V. Palcullo, Donna May C. Rivera, Via Marie F. Dumenden, and Ma. Cynthia T. Peleña, Central Philippine University; Joy F. Geromiano and Daryl Superio, Southeast Asian Fisheries Development Center, Philippines
- Grey literature and the role of women pioneers in Qajar era (1910-1920)** 111
Somayeh Sadat Hashemi, Roya Aminimalroaya, Nooshin Hakhamaneshi, and Ehsan Rezvani; National Library and Archives of Iran

III. Exploring Next Generation Grey – Panel Survey

- Data from "Exploring Next Generation Grey" including Questionnaire and Results** 117
Joachim Schöpfel, University of Lille, France;
Dominic Farace, GreyNet International;
Silvia Giannini and Anna Molino, Institute of Information Science and Technologies, ISTI-CNR, Italy;
Veronika Potočnik, National and University Library, Slovenia;
Dobrica Savić, Nuclear Information Section, NIS-IAEA, United Nations;
David Baxter, University of Alberta, Canada;
Tomas A. Lipinski, School of Information Studies, University of Wisconsin—Milwaukee, United States

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.

The Impact of Digital Transformation on the Sustainability of Grey Literature

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

Abstract

Digital transformation makes an impact on organizations and businesses affecting many of their activities, either in a positive or a negative way. Once an organization starts on the road of digital transformation the impact is always considerable and long-lasting. Due to the specific nature of digital transformation to be able to impact whole industries, even if some organizations decide not to start with digital transformation, they may experience an indirect impact. The area of grey literature management is one of those activities that will have to deal with the indirect impact imposed by the digital transformation of host organizations and related activities. The nature of information work in general, the workforce, and the workplace are undergoing major changes. The same is the case with grey literature. Once it is severely impacted, its long-term sustainability might come into question. This paper deals with the sustainability of three main aspects of grey literature management — the nature of grey literature and the related work, the workforce, and the workplace. To provide sustainability of grey management some specific preconditions need to be met. For example, availability relates to long term preservation, and it includes physical and electronic storage. Also, efficient search and retrieval, together with format recognition, which is directly related to usability, are important preconditions of sustainability. All the preconditions mentioned are challenging tasks in the long run due to fast developments and frequent changes of IT systems, formats, standards, and protocols. All systems need to be fully operational and well maintained, which requires periodic updates, changes, and if needed, complete replacements. This is especially the case with outdated and propitiatory file formats that might become unusable over time. Despite all challenges, digital transformation provides the opportunity to enhance the management of grey literature, increase its value and importance, and improve its sustainability.

Keywords: digital transformation, grey literature, sustainability

Introduction

Our human nature makes us very curious about the world around us. However, whenever we face something new, like a new event, new gadget, new book or article, or a new topic such as this one, we immediately try to evaluate it to figure out what it is all about. More importantly, we view it from our own perspective and wonder why we should care about it or spend our time and energy on it.

The reasons why we should care about the impact of digital transformation on the sustainability of grey literature are numerous. Here are some of the more important ones:

- We are flooded with tons of information and documentation¹.
- It has become hard to keep track of various formats, in particular the grey literature formats. For example, the GreyNet website lists over 150 document types including databases, data sets, datasheets, data papers, satellite data, and product data².
- We are deeply immersed in the digital transformation of current business models.
- Digital transformation impacts the way we conduct our business.

1 The information overload phenomenon has been known by many different names, including: information overabundance, infobesity, infoglut, data smog, information pollution, information fatigue, social media fatigue, social media overload, information anxiety, library anxiety, infostress, infoxication, reading overload, communication overload, cognitive overload, information violence, and information assault (D. Bawden 2020).

2 <https://bit.ly/3naG2E1>

- Digital transformation impacts the grey literature work, the immediate workplace, and the workforce involved.
- Lack of proper care for grey literature brings loss of data and information and loss of knowledge.
- We need to ensure the long-term availability of grey literature, the possibility for its retrieval, use, continuous value, and operability.

This paper looks at the concept of digital transformation and sustainability. It continues with the impact of digital transformation on the grey literature, in particular on the nature of grey literature work, on the grey literature workforce, the leadership, and the impact on the grey literature workplace.

At its end, the paper lists some of the grey literature sustainability principles and offers some conclusions.

Digital transformation

The first concept that needs to be defined is digital transformation. We are all immersed in it, but hardly ever have enough time to define and comprehend it sufficiently enough either from a theoretical or practical aspect.

A brief definition regards digital transformation as a process that leverages modern information technology and brings a large-scale change to business models, processes, and customer experiences in order to create value (i-SCOOP 2019).

The main characteristics of digital transformation are:

- Creation of new business models using modern information technology (IT) and artificial intelligence (AI).
- Streamlining of production processes.
- Focus on customers' experience.
- Leveraging of existing knowledge.
- Change of organizational culture.
- Use of modern data and information management analytical tools.
- Emphasis on "values" and not "activities".

Sustainability

Our next concept that needs some clarification is sustainability.

It is the ability to continue at a particular level for a period of time. It also covers the present time by meeting the needs of the present without compromising the ability of future generations to meet theirs (Mollenkanp 2021).

Another, and probably the most well-known definition of sustainability is the result of four years' work by the Brundtland Commission. According to their report "Our Common Future", sustainable development is:

Development that meets the needs of the present without compromising the ability of future generations to meet their own needs (Brundtland Commission 1987).

Sustainability has three main pillars: social, environmental, and economic. These three pillars are sometimes referred to as people, planet, and profits.

Sustainability helps meet the future needs of both people and businesses by maintaining the required resources. It deals with risk management, saves cost, and potentially drives innovation. It is usually associated with higher quality and in business, it refers to maintaining or sustaining profitability through the use of its assets.

Economic sustainability is a main concern regarding the grey literature sustainability. It is important to decide what sustainability goals will the organization implement, how will success be measured, what are the relevant standards for providing assurance on

sustainability, and how will they be implemented. Of great importance is also the role of leadership in sustainability and how governance will function in the long run.

Grey literature

For the readers, the term grey literature is well known and fully understood. Therefore, I will offer just a very short reminder definition.

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).

This definition considers all major elements of the grey literature concept. Namely, long term preservation, sustainability, usability, and value, while acknowledging the lack of a traditional peer-review process of regular ‘white’ literature.

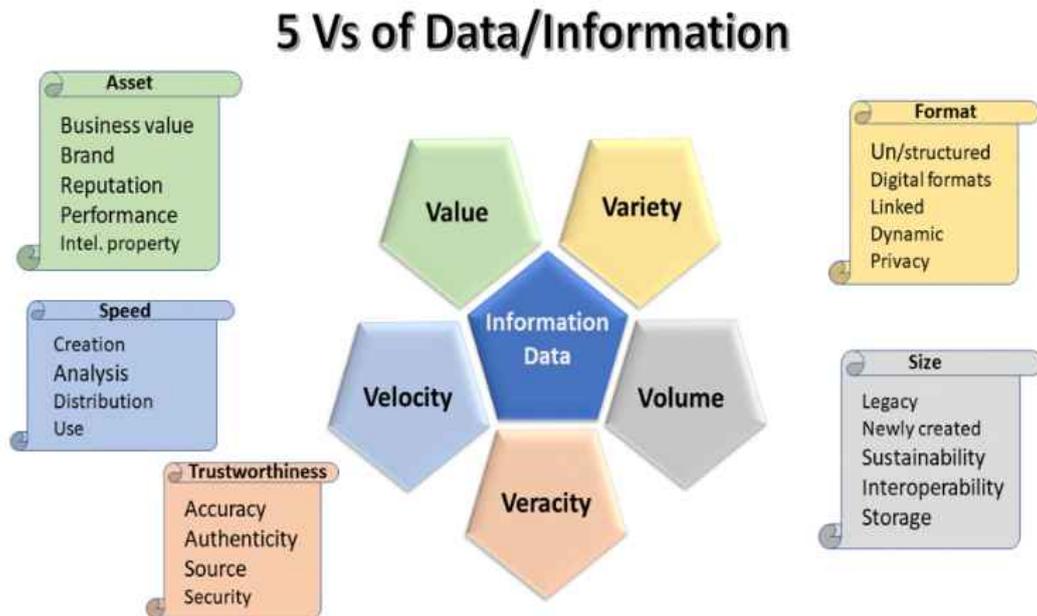
There are many new sources of data, such as the Internet of Things (IoT), Machine to Machine communication (M2M), self-driven cars, robots, sensors, security systems, surveillance cameras, and many other systems or apps using AI and machine learning.

The estimated number of currently connected electronic devices creating specific data varies by billions. Data produced by these devices is highly contextual and software dependent, making it hard to collect and process, and even harder to make sense of and preserve for future use.

Impact of digital transformation on the nature of grey literature work

Digital transformation impacts grey literature sustainability by changing the nature of grey literature work, placing new demands on the grey literature workforce, and also by changing the workplace.

Let’s look at the changes brought to the nature of grey literature work. They are known as the **5 Vs** and are important in any analysis of information and data but also in knowledge analysis.



Drawing 1: Figure 1: 5 Vs of data/information

The 5 Vs represent the following data and information facets:

- **Variety** - new digital formats, unstructured, interlinked, dynamic, privacy issues.
- **Volume or size** - with legacy issues, interoperability, storage, possibility to manage it.
- **Veracity or trustworthiness** - including accuracy, authenticity, source control, and security.
- **Velocity or speed of creation** - analysis, distribution, use, and usefulness.
- **Value** - everyone's concern about information assets that cover business value, branding, reputation, performance, and intellectual property.

Impact of digital transformation on the grey literature workforce

Impact on the grey literature workforce and the roles of grey literature professionals is very sensitive. For example, some previous tasks and functions might be completely transformed, so functions such as search and identify could be replaced by a cognitive search, evaluation, and review by AI algorithms. Collection and processing might change into harvesting. Sharing and promotion might take a road of open access, while long-term maintenance and preservation might come against a stumbling block called organizational policies.

Requirements for new competences

To provide sustainable services and long-term access to scientific grey literature, professionals working on it will need a new set of competencies. Almost all the ones listed below are important. However, some might take more, while some might take less time and energy. Digital literacy and technical knowledge will become paramount, so lifelong microlearning and personal development become a must. Emotional intelligence, social skills, cultural and other diversity will need to be accompanied by a high level of digital ethics. It is assumed that this list of competency requirements will continue to grow.

Impact on leadership

As part of the general workforce, the impact on leadership is already substantial. Leaders are centre-stage for digital transformation. The very nature of leaders' responsibilities undergoes substantial changes because in their new roles they are expected to navigate through the multitude of opportunities, identify the right path for their business and drive adequate and timely change. To that end, digital transformation also brings a lot of uncertainty (Bongiorno 2018).

Traditional approaches for managing business and IT no longer apply. New approaches are required, but they are emerging slowly. Any digital transformation effort will not simply affect IT function alone. It will impact an entire organization, its customers and partners. Risk management is one of the characteristics a new breed of leaders will need to develop since radical changes bring a possibility of "getting it wrong" and paying a high price for it.

Customer focus, good strategy and excellent communication, team building, quick learners, agile management and employee empowerment are just some of the bigger roles leaders will have to assume.

Impact of digital transformation on the grey literature workplace

The impact of digital transformation on the grey literature workplace and long-term sustainability is especially complex. Some of the major stumbling blocks include the old style of management, strict hierarchy, lack of flexibility, complex structures, legacy solutions, and many others. The way out of this, and the road to sustainable scientific grey literature, is through the use of new IT tools, digital dexterity, strong digital culture, removal of information silos, agility, team building, and remote work.

Grey literature sustainability principles

In order for something to be regarded as sustainable, grey literature included, the following criteria need to be met. It needs to be available for a long period of time, be operational and retrievable in the future, and remain usable and valuable.

Conclusions

This paper covered several interrelated concepts, all of them important, and worthy of further analysis and review. However, the most imperative conclusions worth mentioning are the following:

Digital transformation is already with us, it impacts all aspects of our work and it changes the way we create, disseminate, use, and preserve grey literature.

The **changing nature of grey literature** is seen by the increase in GL types and volume, the speed of its creation, the trustworthiness, and its value.

Grey literature sustainability requires that the collections are easily available, retrievable, and usable, as well as remaining valuable and operational.

Grey literature professionals need to develop new digital mindsets so that they can continue to be contributing and respected staff members of future organizations.

Grey literature leaders should improve organization competitiveness and productivity to achieve better results and high-quality services by leveraging IT technology, focusing on customers, empowering employees, and using analytics.

The **grey literature workplace** has already experienced many changes with the introduction of a 'new normal' brought about by the need to adapt to the COVID-19 pandemic by working from home, and still much more change is to come.

At the end of the day, we should remember the famous quote by Charles Darwin, *"it is not the strongest of the species that survive, nor the most intelligent, but the one most responsive to change."*

References

Bawden D., Robinson L., 2020. Information Overload: An Overview. Oxford University Press (OUP). <https://bit.ly/3n2Yvmc>

Bongiorno G., Rizzo D., Vaia G., 2018. CIOs and the Digital Transformation: A New Leadership Role. Springer.

Brundtland Commission, 1987. "Our Common Future: Report of the World Commission on Environment and Development". UN Documents. <http://www.un-documents.net/ocf-02.htm>

i-SCOOP, 2019. Digitization, digitalization and digital transformation: the differences. <https://bit.ly/2r7YPpk>

Mollenkamp D.T., 2021. Sustainability. Investopedia. <https://bit.ly/3HKAmZr>

Savić, D., 2017. Rethinking the Role of GL in the Fourth Industrial Revolution. 10th 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.

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

Grey Literature in Open Repositories: New Insights and New Issues

Joachim Schöpfel and Eric Kergosien, University of Lille- GERiiCO, France
Hélène Prost, CNRS - GERiiCO, France; Florence Thiault, University of Rennes 2, France

Abstract

HAL is the national open repository for documents and data from French scientists. The paper analyses how grey literature is represented on HAL. It presents original empirical results from a follow-up study to former research, based on the scientometric assessment of deposits on the French national HAL repository by more than 1,200 research laboratories. These laboratories are affiliated to ten large French research universities and cover the whole range of scientific domains. We assessed the distribution of document types, the degree of openness, the use of open licenses and the attribution of a DOI. The results are discussed under three aspects: the development since 2019, reuse rights and identifiers.

Keywords: Open science, open access, open repositories, research laboratories, grey literature

Introduction

Open repositories and in particular institutional repositories have been described as “home for grey literature” (Luzi, 2010). They are complementary channels for the dissemination of academic research output insofar they contain other versions of published articles, chapters or books (preprints) and, moreover, documents that have (and probably will) not be published via the usual channels of academic publishing. Part of the open science movement, open repositories are a major factor for the global and free dissemination of research results. Following the EU Open Science Monitor¹, they represent about 25% of the published research, with large differences between countries and disciplines.

Since more than ten years, we assess the deposit of grey literature in open repositories, especially in France, to provide empirical elements for a better understanding of this part of scientific communication and to make recommendations for a better findability and accessibility of grey resources in repositories (see Schöpfel & Prost, 2010, 2014; Schöpfel et al., 2018, 2019, 2020; Stock & Schöpfel, 2009). Among our main findings:

- All open repositories contain grey literature².
- About one third of all deposits is grey literature.
- The part of grey literature is increasing over the years.
- The accessibility (degree of openness) of grey literature is higher than of commercial academic publications.
- The accessibility of grey literature varies between different repositories and between different document types, theses, reports and working papers being generally more open than conference papers.

In 2020, we presented original empirical results of 973,968 HAL deposits (30% of the total HAL content³) of more than 1,200 research laboratories from the ten most distinguished French research universities, including the University of Paris-Saclay ranked #13 by the

¹ EU Open Science Monitor https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/open-science-monitor/trends-open-access-publications_en

² Grey literature has been defined in various manner (Farace & Schöpfel, 2010; Schöpfel, 2011). Usually, grey literature is described as “unpublished”, “not peer reviewed” and “not in databases” and meaning most of the time reports and conference papers (Schöpfel & Prost, 2020).

³ HAL is the national open repository for French public research, with more than 2.6 deposits in August 2021 <https://hal.archives-ouvertes.fr/>

2021 Shanghai ARWU⁴. These laboratories cover the whole range of scientific disciplines, including medicine, law, economics and management. Based on this research, our paper contributes to a deeper understanding of the place of grey literature in open repositories through three complementary studies:

1. The evolution of the deposit of grey literature: is the part of one third stable over the time? How does the composition of grey literature change with the years?
2. The licensing of grey literature: what are the conditions of reuse of grey literature? Is open access to grey literature more “gratis” or more “libre”?
3. The existence of DOI for conference papers: which is the part of commercial publishing of conference papers?

In 2020, the major French research organisation CNRS decided the mandatory use of HAL for the reporting and assessment of the performance of the CNRS research laboratories and individual researchers. For this reason, the analysis of the 2021 deposits provides an exhaustive and reliable photography of the academic output and of the part of the grey literature.

Methodology

This paper presents results from an ongoing research project on open access strategies of French research laboratories⁵; it is a follow-up of two recent studies and employs the same methodology (Schöpfel et al., 2019, 2020). The sample consists of 1,213 research laboratories from the ten universities part of the French excellence initiative (IDEX). We used the laboratories’ identifiers in the HAL repository for the API search of each laboratory’s deposits (February 2021). The results were analysed based on the laboratory and deposit metadata (domain and discipline, university, deposit type, resource category). The API search retrieved 26 different HAL resource categories⁶; we merged and described them in the same way as in 2020 in order to simplify the description of the results (see Annex 1). In addition to the former assessment, we also analysed the availability of a DOI and the use of an open license.

Results

The part of grey literature

The API query retrieved 1,035,612 deposits which have been authored or co-authored by scientists affiliated to one of the 1,213 research laboratories of our sample. From all these items, 33.4% fall under the category of grey literature (figure 1).

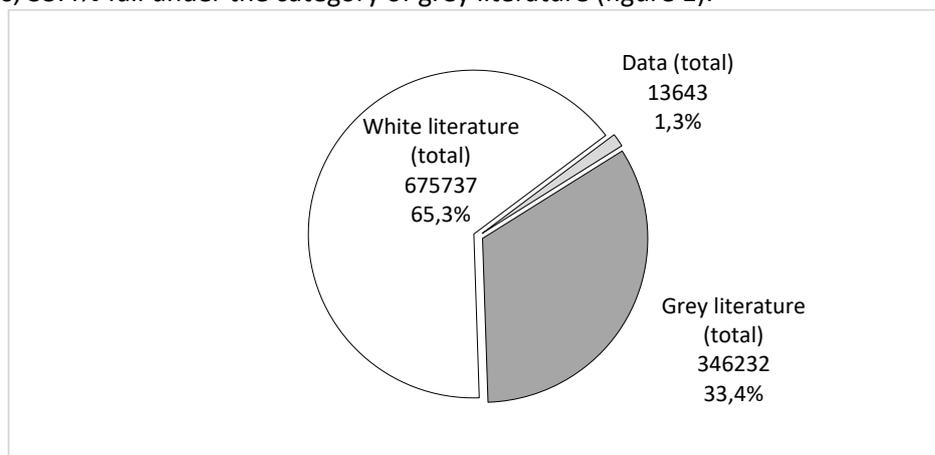


Figure 1. Part of grey literature (N=1,035,612 deposits)

⁴ Academic Ranking of World Universities <https://www.shanghairanking.com/>

⁵ Project HAL/LO <http://gis-reseau-urfist.fr/hal-lo-valorisation-sur-hal-de-la-production-des-laboratoires-dans-lenvironnement-de-la-science-ouverte/>

⁶ These categories may change in the future; in April 2021, a working group recommended a revision of the actual typology, with some new document types, such as data papers and data management plans; cf.

<https://www.ccsd.cnrs.fr/2021/04/evolution-de-la-typologie-des-documents-dans-hal-les-resultats-du-groupe-de-travail/>

Most of the grey literature consists of conference papers (70%), followed by PhD theses (11%) and working papers or preprints (8%). The different types of reports (project reports, activity or annual reports, short reports and report chapters) represent 5% while posters represent 4%. Other resource types are less important, such as BA and Master dissertations, habilitation theses or lectures, totalling together 2% (figure 2).

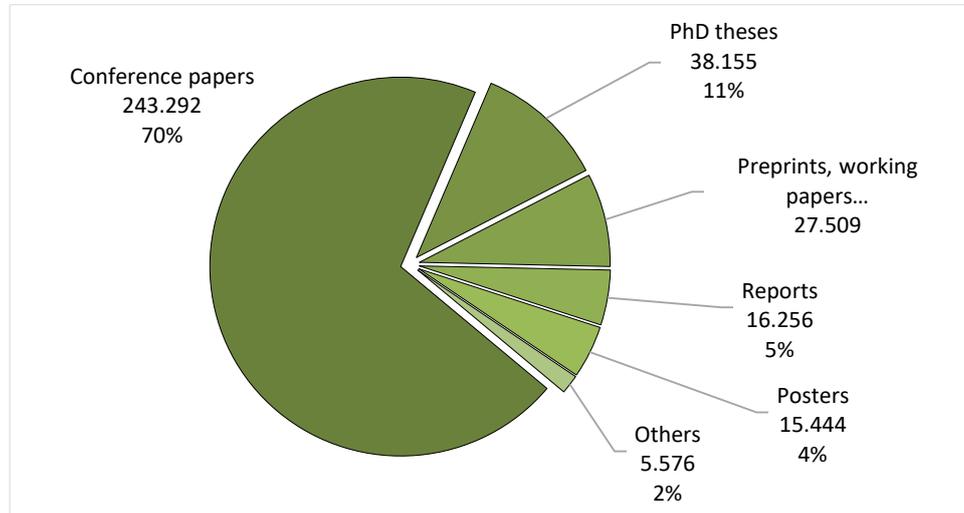


Figure 2. Types of grey literature (N=346,232 deposits)

Figure 1 shows that 1% of the deposits are neither white nor grey literature but datasets. HAL was designed as a document server but contains since 2010 the MédiHAL portal with deposits of visual and sound data (still images, videos and sounds), produced within the framework of scientific research. Also, there are some software deposits (codes) due to the partnership between HAL and the international Software Heritage project.

Degree of openness

The HAL repository contains deposits of full text and files as well as records, i.e., metadata without documents or data files. The part of deposits of full text in our sample is 32.3% which is similar to the overall percentage in HAL (32.9% in August 2021). This part of freely and openly available research output can be interpreted as degree of openness. If all deposits would consist of metadata and data (documents), this degree of openness would be 100%.

Figure 3 shows that the part of items with document and/or data files is significantly higher for grey literature (37.6%) than for white literature (28.2%).

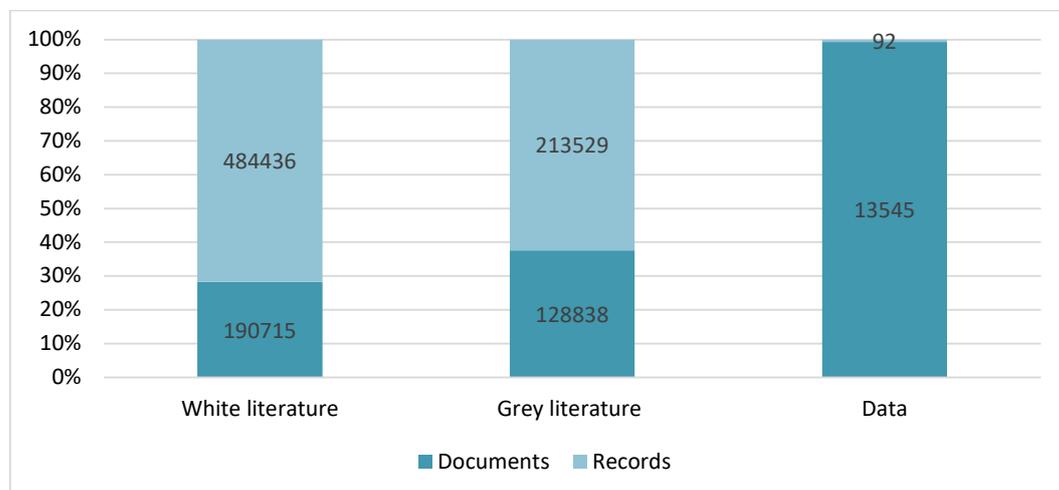


Figure 3. Deposits with and without document and data files (N=1,031,155 deposits, without annex files)

The differences between the document types are important. Figure 4 shows the part of full text for the different types of grey literature.

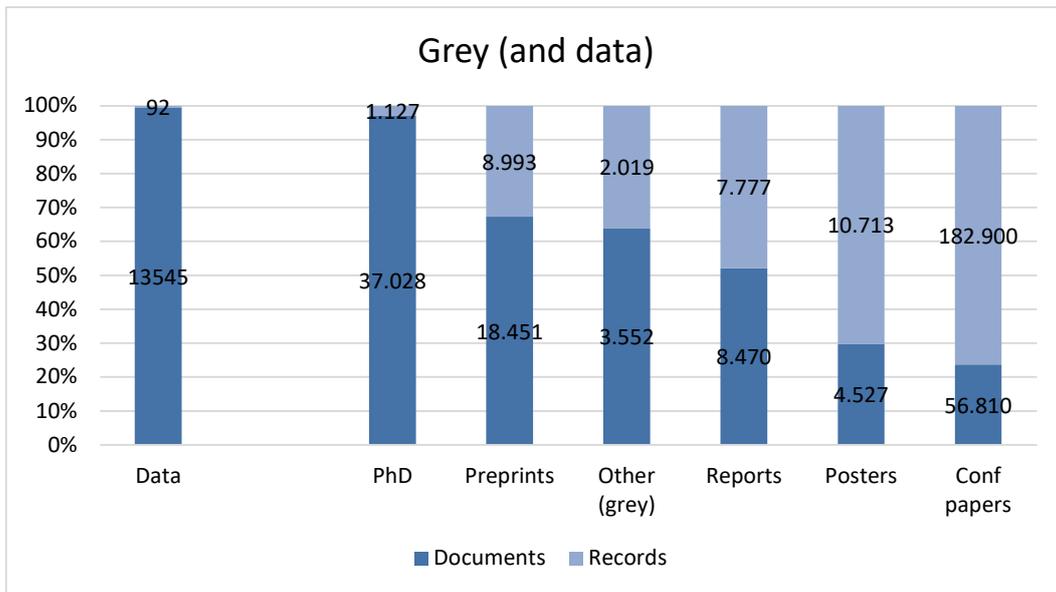


Figure 4. Grey literature deposits with and without document files (N=346,232 deposits)

The degree of openness of grey deposits is generally well above the average percentage, between 50% and near to 100%. The explanation of the exceptional part of openly available PhD theses (97%) is that the self-archiving of a PhD thesis on the HAL platform requires systematically the deposit of the text file.

The open part of the conference contributions, papers and posters, are lower, with 29% open posters and 23% open papers. This lower degree of openness is similar to the percentage of full text deposits of articles, books, chapters etc. (figure 5).

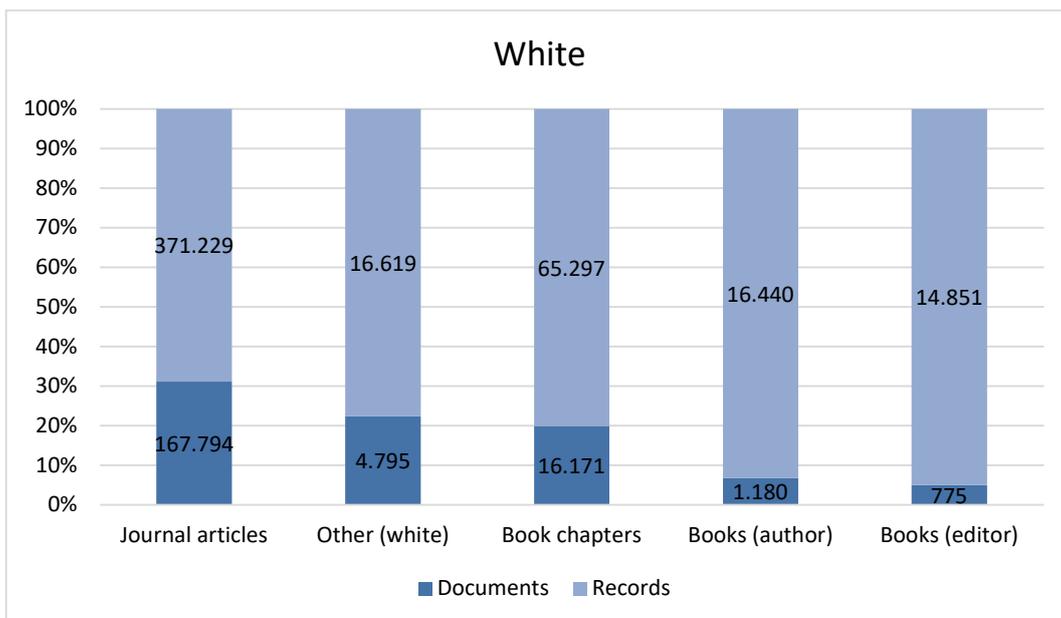


Figure 5. White literature deposits with and without document files (N=657,737 deposits)

31% of the journal articles are openly available on the HAL platform; the degree of openness of chapters, edited and authored books and other white resources is even lower.

Two other observations may be interesting. First, the part of grey literature of all deposits with full text is 40% which is higher than the overall part of grey literature (34%). Second, figures 3 and 4 include the degree of openness of datasets which is exceptionally high (99%); nearly all datasets have been deposited with the data files. Again, the reason is rather simple: MédiHAL, the data portal of HAL, requires the deposit of the data files for each data deposit. Only some software records have been created without the code files.

Disciplinary differences

Each research laboratory has been indexed with a large scientific domain and with a more specific research discipline. In the following, we present the analysis of the HAL deposits regarding four large scientific domains, i.e., science and technology (SciTech), life and medical sciences (BioMed), social sciences and humanities (SSH), and law, economics and management (Law, Econ). Figure 6 shows that there are significant differences between the four domains.

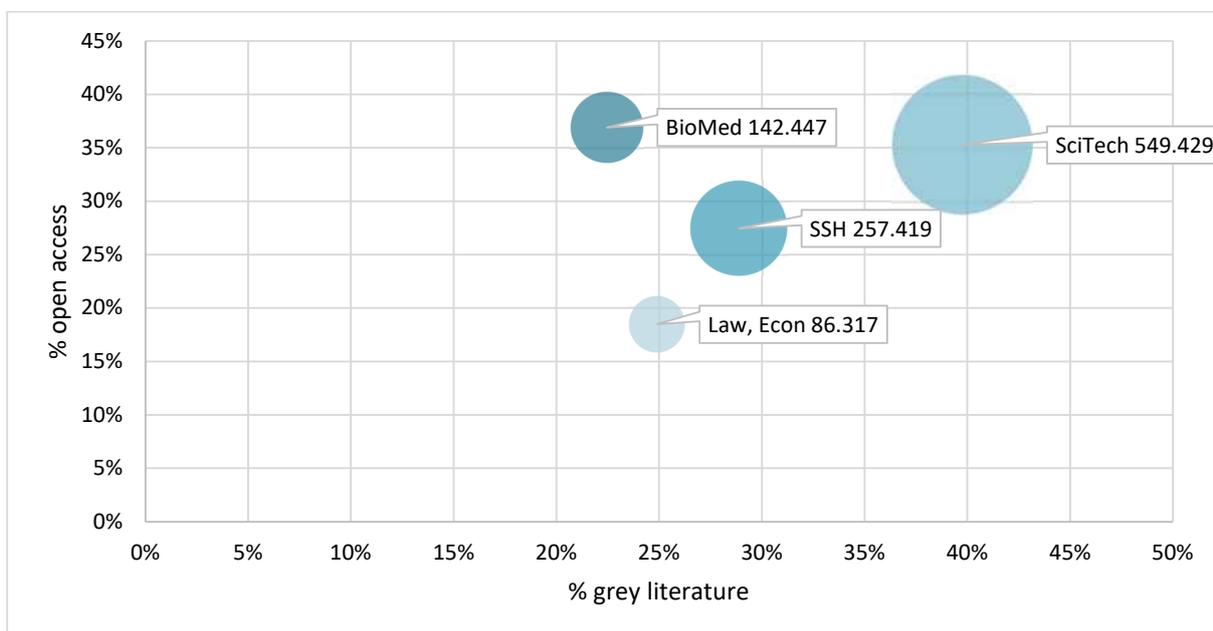


Figure 6. Degree of openness and part of grey literature in four scientific domains (N=1,035,612 deposits)

The deposits of the laboratories in life and medical sciences have the highest degree of openness (37%), followed by those in science and technology (35%). The same indicator is lower in social sciences and humanities (27%) and in law, economics and management (18%).

On the other hand, the laboratories in science and technology have the highest part of grey literature (40%), followed by those in social sciences and humanities (29%), law, economics and management (25%) and life and medical sciences (22%).

Figure 7 presents additional results with a more detailed distinction of ten scientific disciplines; the size of the bubbles represents the number of deposits (for the complete figures, see Annex 2).

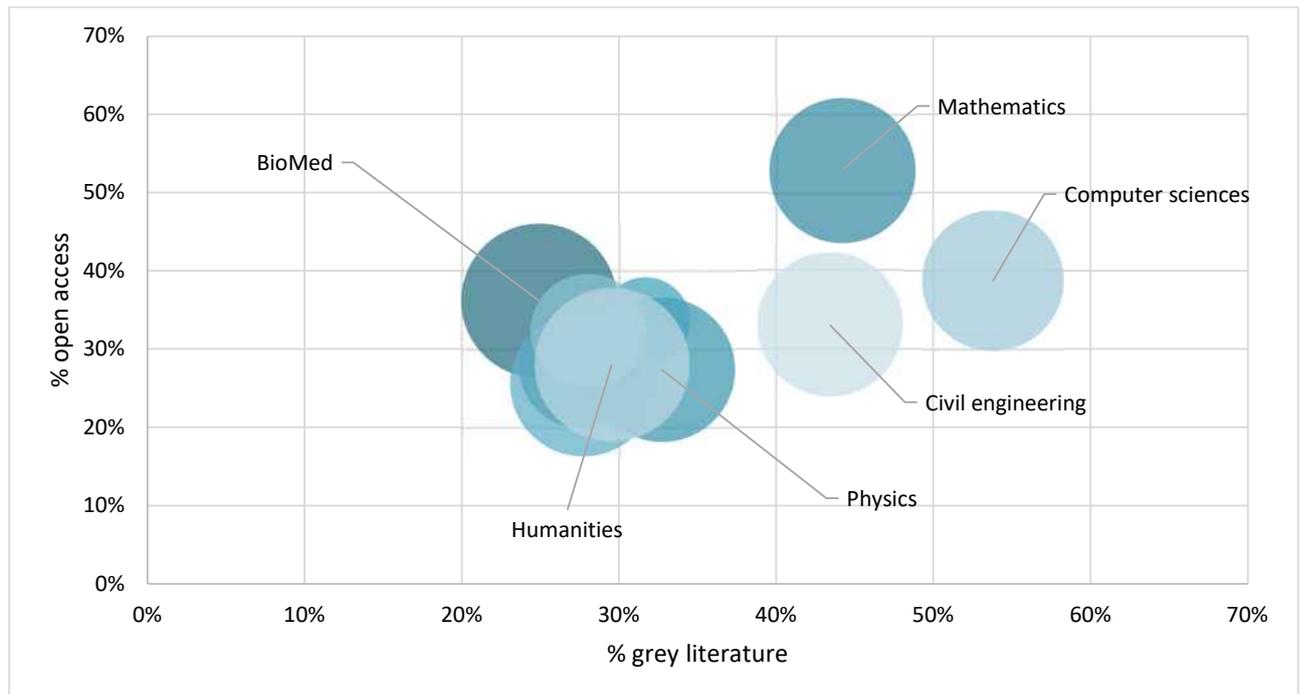


Figure 7. Degree of openness and part of grey literature per discipline (N=1,035,612 deposits)

For most disciplines, the grey part of the HAL deposits represents more or less 30%, varying from 25% (life and medical sciences) to 33% (physics). Three disciplines have significantly more grey literature: civil engineering (43%), mathematics (44%) and computer sciences (54%). Conference papers represent most of this grey literature.

Regarding open access, the part of openly accessible documents is about 30% for most disciplines, varying from 26% (social sciences) to 36% (life and medical sciences). The degree of openness is higher for computer sciences (39%) and particularly for mathematics (53%).

The mapping of document types against scientific disciplines reveals significant differences regarding open access (degree of openness) and relative importance (part of all deposits) (see table in annex 3). Some observations:

- The only grey document type that is really important are conference papers, especially in life and medical sciences (16%), mathematics (9%) and computer sciences (7%). Compared to the overall number of deposits, other grey resources like reports, preprints or working papers are much less important (even if they may contain unique and significant content).
- For conference papers, the degree of openness is higher in mathematics (39%), humanities (37%) and computer sciences (33%) than in physics (13%) or social sciences (12%).
- Regarding reports, mathematics, computer sciences and civil engineering are the disciplines with the highest degree of openness (>80%).
- Regarding working papers and preprints, humanities, mathematics and computer sciences are the disciplines with the highest degree of openness (>70%).

DOI

A digital object identifier (DOI) is a persistent identifier used to identify objects uniquely. The DOI system has been introduced in 1998 and standardized in 2012 (ISO 26324). Following the DOI website, approximately 257 million DOI names have been assigned to date⁷. While the initial focus on entities was documents/media (e.g., articles and data

⁷ DOI <https://www.doi.org/factsheets/DOIKeyFacts.html>

sets), the DOI system is now moving into parties and licences and extending to other sectors.

The initial initiative was taken by the three major international publishing trade associations in order to develop infrastructure for digital publishing. Today, a large part of DOIs is still attributed to articles, book chapters etc. by stakeholders of the traditional publishing economy.

The analysis of the HAL deposits provides the opportunity to assess the part of DOIs assigned to non-conventional literature, such as dissertations, reports, working papers and communications.

Of 1,179,145 analysed items, 486,474 have a DOI (41%). 90.5% of these deposits with DOI are white literature, mainly journal articles (88%) and book chapters (2%). 9.5% of the deposits are grey literature, mainly communications (9%), posters (0.2%) and preprints (0.3%).

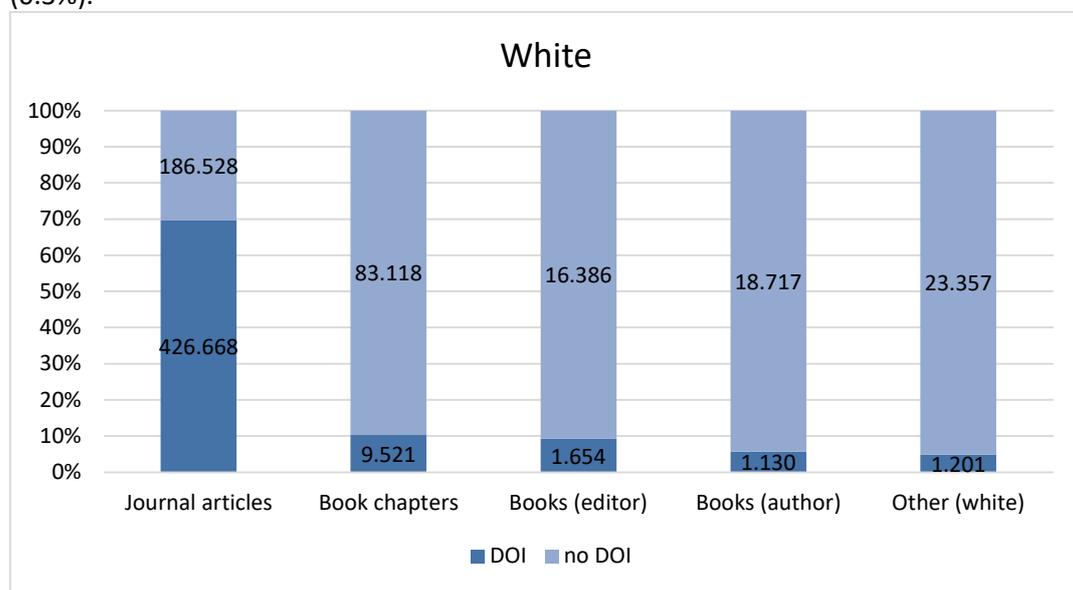


Figure 8. White literature with DOI (N=768,280 deposits)

While 70% of all deposited articles have a DOI (the percentage for book chapters is 10%, for edited books 9%) (figure 8), the part of communications with DOI is only 15% (posters 5%, preprints 4%) (figure 9).

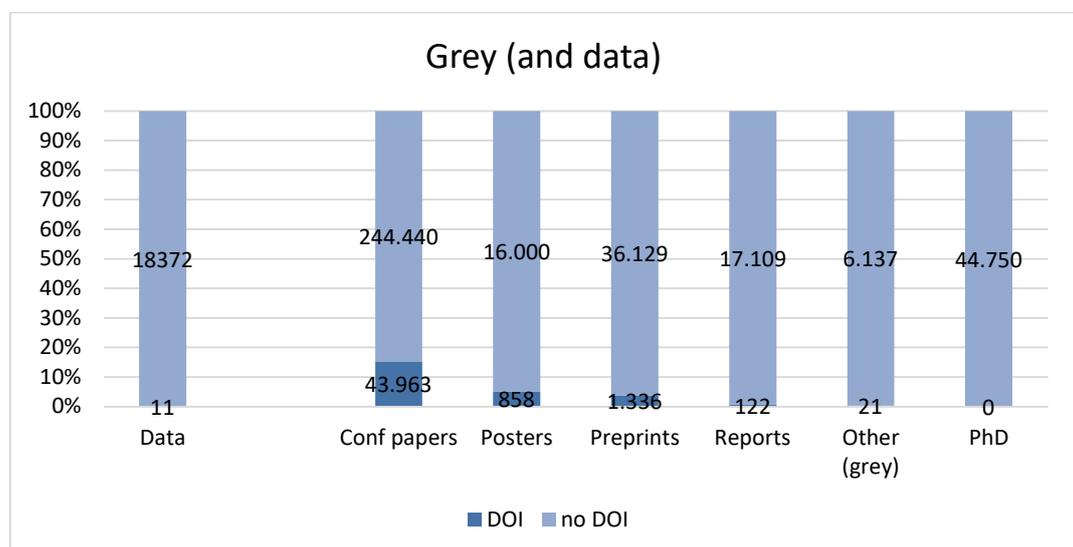


Figure 9. Grey literature with DOI (N=410,865 deposits)

Licensing

By default, the deposits of documents on HAL are published under the French IP law. However, HAL offers the possibility to publish with a license, under copyright or in the public domain, with ten different options. We analysed the legal regime of 351,672 published documents on HAL (table 1).

Legal conditions	Documents	%
By default (French IP law)	298 906	85,0%
Copyright	9 972	2,8%
CC-BY	19 309	5,5%
CC-BY-SA	2 178	0,6%
CC-BY-NC	4 003	1,1%
CC-BY-NC-SA	2 186	0,6%
CC-BY-ND	1 078	0,3%
CC-BY-NC-ND	10 394	3,0%
CC0 Public Domain Zero	82	0,0%
NC Public Domain Mark	29	0,0%
Etalab Open License	99	0,0%
Public domain	3 436	1,0%

Table 1. Legal regime of publishing (N=351,672)

85% of the documents are published under the French IP law, by default. 11% of the documents are published with a Creative Commons⁸ or Etalab⁹ open license, 1% are published in the public domain, with a CC Zero Public Domain Dedication (CC0), with a NC Public Domain Mark or with a simple public domain indication (figure 10).

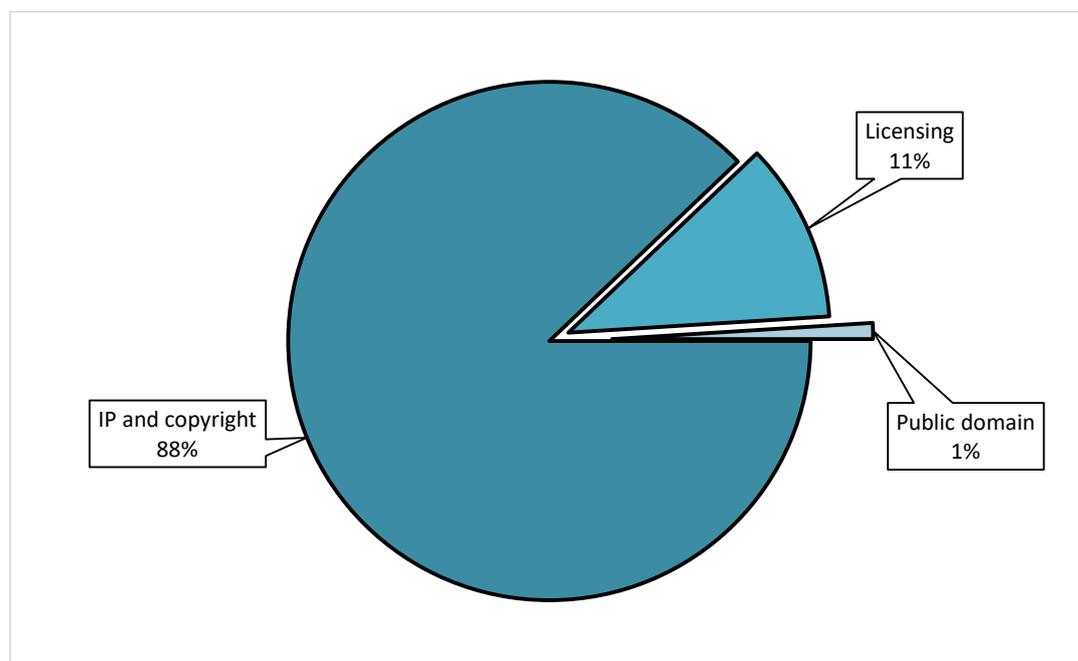


Figure 10. Legal regime of publishing (N=351,672)

39,148 documents are published with a Creative Commons license (11%, without CC0). Nearly half of these documents are published with a (liberal) CC-BY Attribution license,

⁸ Creative Commons <https://creativecommons.org/>

⁹ Etalab is the French public agency in charge of the open data strategy <https://www.etalab.gouv.fr/>

26% are published with a (more restrictive) CC-BY-NC-ND license which doesn't permit any commercial use and any derivatives or adaptations of the work (figure 11).

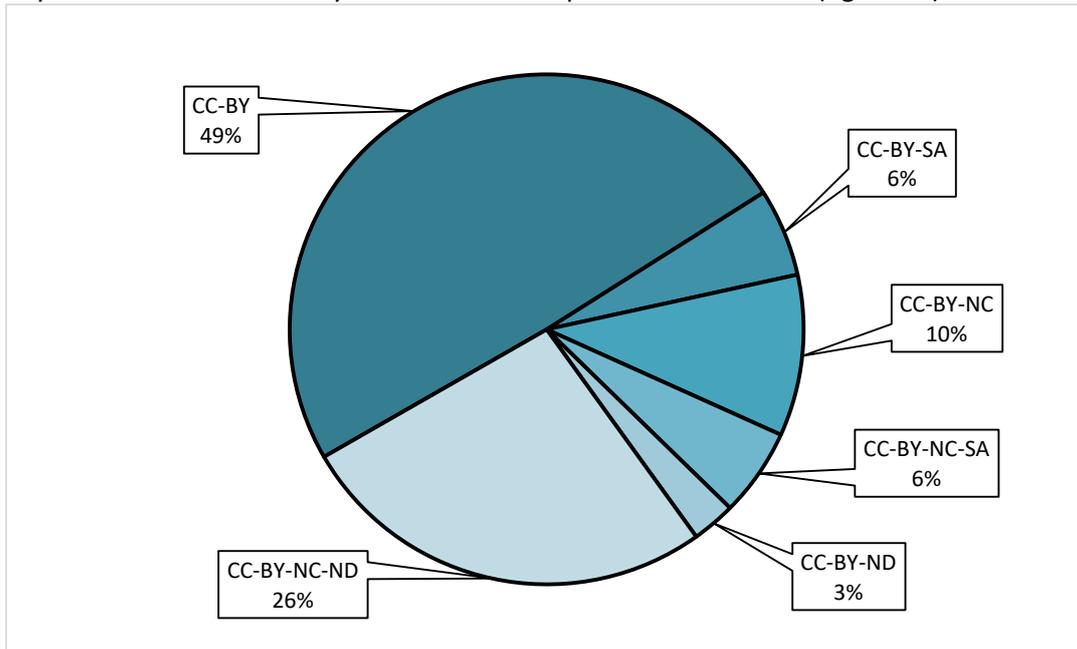


Figure 11. Creative Commons licenses (N=39,148)

For a small subsample of 22,508 deposited files from 70 laboratories we analysed the distribution of open licenses and document types. The results can be resumed as follows:

	Grey documents	White documents
CC-BY	1114	3676
CC-BY-SA	377	262
CC-BY-NC	205	501
CC-BY-NC-SA	271	352
CC-BY-ND	81	454
CC-BY-NC-ND	941	2316

Table 2. Publishing with open licenses (N=10,550)

- Grey items are less often published under an open license than articles, book chapters and other “white” documents. In fact, open licenses are most often applied to the sharing of research data and other materials, like images, AV files or software.
- Regarding different grey items, especially PhD theses are very seldom published under an open license, less often than posters, reports and other, miscellaneous documents.
- The most important licenses applied to grey literature are the liberal CC BY (Attribution) license (37%) and the restrictive CC BY-NC-ND (Attribution-NonCommercial-NoDerivs) license (31%). The main difference with articles, books and so on is that the part of the Wikipedia CC BY-SA (Attribution-ShareAlike) and the more restrictive CC BY-NC-SA (Attribution-NonCommercial-ShareAlike) licenses is more important for the publishing of grey literature (together 21%, against 8% for white items) (table 2).
- Publishing in open domain, for instance with a CC0 (Creative Commons Public Domain Dedication), plays nearly no role for grey items, similar to articles, books and so on.

Discussion

Changes

Since 2020, the number of grey items in our sample of more than 1,200 research laboratories increased by 2.5% (= 8,491 new deposits). This is less than the overall growth of all deposits (6.3%) and less than the articles, book chapters etc. (white literature) of which the number has risen by 8.1%. As a result, the part of grey literature decreased from 34.7% in 2020 to 33.4% in 2021. This is a general evolution in all domains and disciplines, even if in some disciplines (especially in civil engineering, computer sciences and physics) the observed decrease is a little bit more important than in others.

There are probably two main reasons for this slight relative decrease. The first is the new open access policy launched by the French research organisation CNRS in 2020; researchers from the CNRS laboratories must deposit their publications on the HAL platform for their individual annual assessment, and as this assessment focusses on published articles, their number and part increased in a significant way. The second reason is the secondary exploitation right introduced by the 2016 Digital Law which foster the deposit of published articles on open repositories and in particular, on HAL, and which doesn't include other document types.

However, this development should not be described as an erosion but is probably more as a kind of ceiling or capping of the part of grey literature at a level of about 33%.

Simultaneously, it can be observed that the part of open access among the grey literature slightly increased from 36.6% in 2020 to 37.6% in 2021, which is consistent with the open access strategy of the French government, the universities and research organisations. The differences of openness between the types of grey literature remained stable and unchanged, except for working papers and preprints where the degree of openness increased by 3%. In a more general way, grey items remain more accessible than articles and books, with one exception, like in the past: the openness of communications, i.e., the accessibility of papers presented at scientific events, remains low. Again, there may be at least two reasons for this lack of openness: first, for one part of the communications, there may be simply no written and publishable papers; second, another part of the communications is not really grey literature, as defined by the Luxemburg, New York or Prague definitions, insofar they are published (and controlled) by traditional academic publishing houses. The attribution of DOIs for 15% of the deposited communications is an indication for this distinction.

Reuse rights

Globally, IP and copyright protection is dominant. Only a small part of documents has been published with open licenses, mostly with the liberal CC-BY license which is the license preferred and fostered by French public research institutions and with the more restrictive CC-BY-NC-ND license. Very few documents have been published in the open domain. Regarding the FAIR principles, this means that the reusability of most of the HAL documents is not very high, in contrast to the research data and other materials deposited on HAL which are more often published under an open license. The results of a small subsample of 70 laboratory collections show that the two Creative Commons ShareAlike licenses are more often used for grey literature than for the sharing of journal articles, book chapters and so on, and that very few PhD theses are published under an open license. However, because of the small size of this subsample we should be careful and avoid generalizing interpretations of these results.

Identifiers

The attribution of DOIs is not limited to documents published by commercial academic publishing houses but in fact, it is. Up to now, very few academic or research institutions assign DOI names to their research output. Perhaps the growing DataCite network will change the game but for the moment, the DataCite DOIs are above all assigned to

research data and software, not to other items, even if (at least in France) academic libraries started a discussion about the attribution of DataCite DOIs to documents like dissertation and theses, reports and so on.

Two aspects should be highlighted: As we have already said before, as long as DOIs are not massively assigned to grey literature like dissertation, reports, working papers, conference papers and so on, these documents will remain largely out of scope of altmetrics tools and open science monitoring, as both are mainly (exclusively) based on DOIs (Schöpfel & Prost 2016 and 2019). “Researchers wishing to assess impact of their own grey literature should endeavour to create DOIs where appropriate so persistent identifiers already used to assess impact by other mediums (Altmetric, n.d.) can be extended into the grey literature realm” (Bickley et al., 2021).

The fact that 15% of the deposited communications on HAL have a DOI can be seen as an indication that these deposits are probably not natively grey but white literature, published in proceedings and journals by publishers like Springer Nature, Elsevier, IEEE and so on. Here, we should be careful with the application of the concept of grey literature to documents which in fact are not (are no longer or have never been) grey items.

We already discussed the “greyness” of electronic theses and dissertations (ETDs) elsewhere (Schöpfel & Rasuli, 2018). In France, ETDs became mandatory in 2016. Since then, each PhD dissertation must be submitted to the French national ETD infrastructure STAR, will be available (if not confidential) via the national ETD portal theses.fr and the European portal DART-Europe and is preserved in a national and sustainable long-term dark archive hosted by the public agency CINES. And if not decided otherwise by the author, it will be accessible on the HAL platform (TEL portal) or on an institutional repository, embargoed or not. Thus, French theses have become significantly since 2016 less grey than before, and that is good news.

Conclusion

Our paper presents empirical results of a follow-up study on the deposits of more than 1,200 French research laboratories on the national open repository HAL. The figures show a significant increase of all deposits since 2020 (+6%), a growth which seems consistent with the global development of scientific performance and with the open science strategy of the French government and institutions.

As in 2020, the part of grey literature is about one third of all deposits, and this part is generally more accessible (with more deposits of document files) than articles and books. The general impression is that of continuity and stability, without significant changes since last year. However, it can be observed that HAL has become (a bit) more open and at the same time, (a little bit) less grey. We have provided some possible reasons for this development, such as the open science strategy of the French government, the secondary exploitation right for French public research and the mandatory open access policy of the CNRS.

In terms of FAIRness, the situation can be described as follows:

FAIR principle	General observation	Grey documents
Findability	Less than half of the deposits have a DOI	Weaker than articles
Accessibility	All deposits are retrievable on HAL with the standard OAI protocol but only for one third the document is accessible	Globally better than for articles and books
Interoperability	n/a	n/a
Reusability	Weak – only 15% items are released with a license	Similar

The comparison between white and grey deposits shows that the two FAIR principles findability and reusability remain a challenge for the production and the publishing of grey literature on HAL, because of the low levels of DOI assignment and of licensing.

As the deposit of publications on HAL becomes more mandatory and thus, more representative for the overall output of French research laboratories, it is possible that we can see here a kind of ceiling of grey literature, on different levels depending on disciplines and institutions. It will be interesting to assess the impact of new forms of scientific communication, such as webinars or preprints, especially in the aftermath of the Coronavirus pandemics. For the moment, we can't see any significant effect on the HAL platform, probably because these new forms are not relevant for the research assessment, career decisions and so on.

The development of conference papers in particular and of all kind of resources related to scientific events in general will be particularly interesting for the understanding of the future of grey literature. They are not only the most important part of the traditional "grey document types" but they are also somehow on the edge, partly grey, and partly controlled by commercial publishing.

The future will also show how the international (European) open science policy in favour of gold open access (article publishing in open access journals) will impact the development of open repositories like HAL. If more and more articles are freely available on journal platforms, maybe that those repositories will become (again) a home for grey literature, not controlled by commercial publishing. Perhaps this will also change our understanding of this part of scientific communication.

Acknowledgments

The study is part of the research project HAL/LO funded by GIS "Reseau URFIST" from 2019-2021 with 10,000 euros. We would like to express our gratefulness to the whole project team for rich discussions and especially to Bernard Jacquemin for data extraction from the HAL platform, and we would also like to thank Bénédicte Kuntziger from the CCSD for helpful advice.

References

- Bickley, M., Kousha, K., & Thelwall, M. (2021). A systematic method for identifying references to academic research in grey literature. *18th International Conference on Scientometrics & Informetrics Proceedings*, 121–132.
- Farace, D., & Schöpfel, J. (Eds.). (2010). *Grey Literature in Library and Information Studies*. Berlin, De Gruyter Saur.
- Luzi, D. (2010). Grey Documents in Open Archives. *The Grey Journal*, 6(3), 137–144.
- Schöpfel, J., & Prost, H. (2010). *Développement et Usage des Archives Ouvertes en France. Rapport. 1e partie : Développement*. Université de Lille 3.
- Schöpfel, J., & Prost, H. (2014). Degrees of Openness. Grey Literature in Institutional Repositories. In *GL16 Sixteenth International Conference on Grey Literature. Grey Literature Lobby: Engines and Requesters for Change*. 8-9 December 2014, Library of Congress, Washington D.C., USA (pp. 75–85). Amsterdam: TextRelease.
- Schöpfel, J., & Prost, H. (2016). Altmetrics and Grey Literature: Perspectives and Challenges. *GL18 Eighteenth International Conference on Grey Literature*. 28-29 November 2016, New York Academy of Medicine, New York NY, USA.
- Schöpfel, J., & Prost, H. (2019). The scope of open science monitoring and grey literature. *12th Conference on Grey Literature and Repositories*, National Library of Technology (NTK), 17 October 2019, Prague, Czech Republic.
- Schöpfel, J., & Prost, H. (2020). How scientific papers mention grey literature: a scientometric study based on Scopus data. *Collection and Curation*, 40(3), 77-82.
- 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.
- Schöpfel, J., Prost, H., Fraisse, A., & Chaudiron, S. (2018). Valoriser les publications d'un laboratoire universitaire dans l'environnement de la science ouverte : Retour d'expérience de la collection GERiICO sur HAL. In *ICOA 2018 3e colloque international sur le libre accès*, 28-30 November 2018, Rabat, Morocco.
- Schöpfel, J., Prost, H., & Ndiaye, E.H.I. (2019). Going Green. Publishing Academic Grey Literature in Laboratory Collections on HAL. In *GL2019, 21st International Conference on Grey Literature. Open Science Encompasses New Forms of Grey Literature*, 22-23 October 2019, TIB, Hannover, Germany.
- 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. *GL2020. 22nd International Conference on Grey Literature, Applications of Grey Literature for Science and Society*, 19 November 2020, Rome, Italy.
- Stock, C., & Schöpfel, J. (2009). Grey literature in French digital repositories: a survey. In *GL10 Tenth International Conference on Grey Literature: Designing the Grey Grid for Information Society*, Amsterdam, 8-9 December 2008 (pp. 39–53). Amsterdam: TextRelease.

Annex

Annex 1 - Resource categories and types

HAL category	Merged category	Resource type
Journal article	Articles	White
Book (author)	Books (author)	White
Book (editor)	Books (editor)	White
Book chapter	Book chapters	White
Other	Other (white)	White
Communication	Conference papers	Grey
Poster	Posters	Grey
PhD thesis	PhD theses	Grey
Report	Reports	Grey
Activity report	Reports	Grey
Short report	Reports	Grey
Report chapter	Reports	Grey
Other report	Reports	Grey
Undefined	Preprints, working papers...	Grey
Habilitation	Other (grey)	Grey
Lecture	Other (grey)	Grey
Master dissertation	Other (grey)	Grey
BA dissertation	Other (grey)	Grey
Lecture note	Other (grey)	Grey
Presentation	Other (grey)	Grey
Patent	Other (grey)	Grey
Image	Data	Data
Map	Data	Data
Software	Data	Data
Audio	Data	Data
Video	Data	Data

The HAL category “other” consists of book reviews, encyclopaedia entries, translations etc., most of them reviewed publications. The HAL category “undefined” consists mainly of working papers, preprints and other, non-reviewed and unpublished documents. Some of the HAL resource categories are specific to a particular, institutional workflow (i.e., ingestion of records from institutional partnerships), such as BA dissertations, lecture notes and report chapters; they represent very small figures (<100) and have been merged (reports) or included in the “other” category.

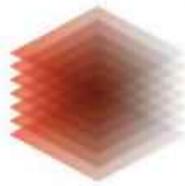
Annex 2 – Disciplinary differences

	Nb of deposits	% grey literature	% open access
BioMed	225 941	25%	36%
Chemistry	135 541	28%	28%
Mathematics	87 176	44%	53%
Physics	197 509	33%	27%
Agronomy, ecology	73 530	32%	34%
Social sciences	199 633	28%	26%
Earth sciences, space	124 108	28%	32%
Computer sciences	185 032	54%	39%
Humanities	223 067	30%	28%
Civil engineering	195 253	43%	33%

Annex 3 – Disciplinary differences and document types

For three grey document types: conference papers, reports and working papers/preprints.

Discipline	Nb of labs	Document type	Deposits	Records	Documents	% deposits	% open access
Agronomy, ecology	21	Conference papers	3 080	2 534	505	4,2%	16,4%
BioMed	356	Conference papers	35 324	25 778	8 816	15,6%	25,0%
Chemistry	38	Conference papers	3 523	2 763	675	2,6%	19,2%
Civil engineering	37	Conference papers	8 933	6 505	2 276	4,6%	25,5%
Computer sciences	24	Conference papers	13 780	9 047	4 526	7,4%	32,8%
Earth sciences, space	10	Conference papers	3 790	2 663	1 047	3,1%	27,6%
Humanities	24	Conference papers	2 242	1 316	821	1,0%	36,6%
Mathematics	12	Conference papers	7 732	4 634	3 027	8,9%	39,1%
Physics	31	Conference papers	8 038	6 877	1 078	4,1%	13,4%
Social sciences	13	Conference papers	1 109	962	131	0,6%	11,8%
Agronomy, ecology	15	Reports	255	155	100	0,3%	39,2%
BioMed	158	Reports	1 717	641	1 076	0,8%	62,7%
Chemistry	14	Reports	81	30	51	0,1%	63,0%
Civil engineering	20	Reports	147	29	118	0,1%	80,3%
Computer sciences	13	Reports	454	65	389	0,2%	85,7%
Earth sciences, space	8	Reports	277	148	129	0,2%	46,6%
Humanities	15	Reports	83	51	32	0,0%	38,6%
Mathematics	8	Reports	364	36	328	0,4%	90,1%
Physics	14	Reports	140	74	66	0,1%	47,1%
Social sciences	11	Reports	126	94	32	0,1%	25,4%
Agronomy, ecology	15	Preprints, working papers	261	180	80	0,4%	30,7%
BioMed	240	Preprints, working papers	3 369	1 082	2 272	1,5%	67,4%
Chemistry	26	Preprints, working papers	575	290	284	0,4%	49,4%
Civil engineering	32	Preprints, working papers	686	221	465	0,4%	67,8%
Computer sciences	20	Preprints, working papers	578	146	429	0,3%	74,2%
Earth sciences, space	10	Preprints, working papers	538	288	250	0,4%	46,5%
Humanities	23	Preprints, working papers	183	33	150	0,1%	82,0%
Mathematics	11	Preprints, working papers	1 003	288	712	1,2%	71,0%
Physics	27	Preprints, working papers	1 091	465	625	0,6%	57,3%
Social sciences	11	Preprints, working papers	219	126	92	0,1%	42,0%



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

Improving guidelines for video abstracts: An analysis of the most popular videos abstracts in the TIB AV Portal

Margret Plank and Jens Kösters,
TIB Leibniz Information Centre for Science and Technology, Germany

Abstract

Today, more and more scientific videos are published online. One visual format that seems particularly suitable for communicating scientific content is the video abstract. This is the 3-5 minutes long moving image equivalent of a written abstract. With this format, scientists have the opportunity to explain the results and background of their concrete research work as well as the methods used, the study results, and possible implications to a potentially larger audience.

Recently, some studies have been published on this subject, in particular with a focus on content analysis for specific domains and classifications of online videos in general. This paper explores the topic of video abstracts and publishing guidelines in order to answer the following question: „Do authors follow the existing guidelines and are the guidelines sufficient for the publication of a scientific work?“. In a literature and web review we looked at the existing publishing guidelines and extracted the major rules. A database from the most viewed thirty-three video abstracts, published on the TIB AV-Portal was created. Each video was analyzed for different criteria such as link to corresponding paper and research data, length, formats etc.. Results indicate that the most common guidelines were followed by the authors, such as max. 4 minutes length, inclusion of additional relevant material such as images, animations, and lab footage as well as good audio quality. However there is still a lot of potential to get more out of video abstracts e.g. adding a title and a link to the corresponding paper and research data.

Keywords *Video Abstracts, Video Platforms, Scientific Videos, Publishing Guidelines, Video Design*

Introduction

The scientific community thrives on the exchange, communication and dissemination of research results, ideas and projects. This takes place to a large extent at conferences, meetings, symposia and workshops, without which the scientific landscape would be inconceivable. Traditionally, the results are published in the form of articles or proceedings, thus documenting the current state of research. The advent of digital media has fundamentally changed information, communication and work behavior in science and society. Communicating research results in the form of or with audiovisual media has become a modern, rapidly growing part of scientific communication (Leon & Bourk, 2020). This has a great potential to communicate scientific findings to a wide audience that would otherwise hardly learn about the valuable research results (Bucher 2020, Leon & Bourk, 2020).

State of the Art Video Abstracts

Scientific videos have gained popularity in the last decade (van Edig, 2016, Ferreira et al., 2021). There are a number of different formats of scientific videos, such as conference recordings, project documentations, video data from qualitative research, and recorded experiments. Studies on the classification of online scientific videos (Boy et al., 2020, Morcillo et al. 2016) also distinguish between presentation, expert, animation, and narrative videos, while others first make a rough distinction between TV and online formats and other subgenres (Garcia-Aviles and Lara, 2020). One visual format that seems

particularly suitable for communicating scientific content is the video abstract (Berkowitz 2013, Ferreira et al 2021) (hereafter abbreviated: VA). This is a 3-5 minute moving image equivalent of a written abstract (Berkowitz, 2013; p.1). This format gives researchers the opportunity to explain the findings and background of their specific research, as well as the methods used, study results, and possible implications to a potentially larger audience (Plank et al 2018). VAs can be embedded in researchers' websites, institutes, or science blogs, shared on social media, and included in lectures. In this way, they develop a visibility that extends into various social and economic spheres. Science journalists, decision makers, and interested lay people can thus learn about the latest results and discussions in science. Similarly, VAs can be helpful in keeping up with the growing body of interdisciplinary research and providing an overview of research outside one's own discipline (van Norden, 2015). Another benefit of VA creation is to revisit one's research findings in a different format. To this end, Dr. Whitesides from Harvard has his students create three-minute summaries of their research findings in VA style (Whitesides, 2011). Spicer (2014) showed in a study using the *New Journal of Physics* (published by IOP Science) that articles accompanied by a VA are more likely to be downloaded than those without. Of the top 25 articles with the highest usage, 36% had an associated VA (Spicer, 2014; p.9). A VA can thus be a useful tool for converting video views into downloads of online articles-especially when published in open-access journals (Watkins, 2016). Moreover, VAs positively influence the citation rates of scientific articles, as Zong et al. (2019) explain in their study - also based on the "*New Journal of Physics*".

Publication and Guidelines

The whole idea behind a video abstract is to publish it online and directly link it with a scientific paper that has been accepted and published in a scientific journal. On the one hand, numerous commercial platforms are available for the publication, such as YouTube or Vimeo. Cell Press¹ was one of the first publishers to recognize the potential of VA and set up its own YouTube Channel as early as 2009. Renowned scientific publishers such as Copernicus Publications, IOP Science, Elsevier, Wiley, and Taylor & Francis have also started to offer the possibility to submit VAs and link them to the respective reference article. The popular transdisciplinary repositories Figshare² and Zenodo³ support the publication of videos as research output. Videos receive a Digital Object Identifier (DOI), descriptive metadata, and author information are available. The open US platform WeShareScience (<http://wesharescience.com>) also provides a place to link, share and discuss VAs. The TIB AV-Portal (<https://av.tib.eu/>) is a reliable, sustainable infrastructure for publishing scientific videos, for example in cooperation with open access publishers such as Copernicus Publications.

Many publishers have published guidelines to help scientists produce their video abstracts (Plank et al 2017). As an example, the Design Rules of the Institute of Physics (IOP)⁴ include the following hints in terms of production:

- A video abstract should not last longer than 4 minutes.
- Inclusion of additional relevant material such as images, animations, and lab footage is strongly encouraged.
- A video abstract must include a soundtrack providing a clear verbal narration of the visual content.

¹ <http://www.cell.com/video>

² <https://figshare.com>

³ <https://zenodo.org>

⁴ <https://iopscience.iop.org/journal/1538-3873/page/video-abstract-guidelines>

Other publishers like Taylor & Francis also give general design recommendations⁵ including: make it short, avoid overload, be natural, be clear and to the point, use images, make sure your audio is clear. Theory Culture & Society⁶ and BMJ⁷ give similar advice. Sage after all complements that a link to the article should be added in the caption⁸. WeShareScience is offering different templates to create a video abstract⁹.

Wiley uses step-by-step guides to explain how to produce and publish video abstracts and makes suggestions where and how to share the video abstracts¹⁰:

- Share it on your social media channels (and ask your department or institution to share it on theirs)
- Include it on your personal website and Kudos publication details
- Include it in a SlideShare about your article
- Use it in conference presentations
- Email it to colleagues and peers
- Send it to your local press office
- Link to it from future grant applications

However in all of these guidelines formal criteria are missing. These should include mention of the authors, the title, a link to the corresponding paper and research data, bibliographical references, licence information and if possible a digital object identifier (DOI).

One portal which is science compliant because the upload process requires formal metadata as well as a licence agreement is the TIB AV-Portal¹¹. Additionally, all videos receive a digital object identifier from DataCite. The aim of the portal is the professional hosting and semantic indexing of scientific videos for research purposes. The portal currently provides more than 36,000 (as of October 2021) quality-checked scientific videos, primarily from technology and the natural sciences, under open access, predominantly Creative Commons licenses. These include computer visualizations, learning videos, simulations, experiments, interviews, lecture and conference recordings, and video abstracts - systematically published, for example, as part of a cooperation with Copernicus Publications. In order to ensure the long-term availability as well as the second-by-second referencing of the videos, they are digitally preserved and receive a Digital Object Identifier (DOI) as well as Media Fragment Identifiers (MFID). In addition, various automated analysis procedures are used in the AV-Portal, through which the videos are indexed on a fine-granular and time-based basis: Temporal Segmentation; Keyframe-based Text Recognition (Video OCR); Speech Recognition for the creation of audio transcripts; Annotation with predefined visual concepts; as well as the semantic analysis and differentiated tagging of the video content with subject-specific keywords. With the help of these analysis methods, information and content can be precisely localized using semantic and explorative search functionalities ("facet search"), videos can be searched for content, and relevant video sections can be retrieved with segment precision. Furthermore, the videos in the portal can be searched across languages via a cross-lingual mapping (German-English) of the semantic knowledge base. The metadata is also available as standard CC0-licensed RDF data for subsequent use.

⁵ <https://authorservices.taylorandfrancis.com/research-impact/creating-a-video-abstract-for-your-research/>

⁶ <https://www.theoryculturesociety.org/video-abstract-guidelines>

⁷ <https://authors.bmj.com/writing-and-formatting/video-abstracts/>

⁸ https://us.sagepub.com/sites/default/files/sage_video_abstract-external_guidelines.pdf

⁹ <https://wesharescience.com/Create-Video-Abstract>

¹⁰ <https://authorservices.wiley.com/author-resources/Journal-Authors/Promotion/video-abstracts.html>

¹¹ tib.av.eu

As mentioned, the portal provides standardized metadata for video abstracts, a license statement, and a DOI. However, it does not ensure that scientists cite their paper and research data directly in the VA. Together with following the guidelines, it is up to the scientists to get the most out of their video abstract.

Study and Method

In our study¹² we wanted to find out whether the most popular video abstracts on the TIB AV portal follow the common design and publishing rules or not. Our study provides a characterization of video abstracts published on our portal in the areas of Science and Technology. We identified video abstracts based on the number of views. A database of the highest ranked thirty-three videos, from four publishers was created. Each video was manually analyzed for different parameters. Based on a literature review (Ferreira M. et al. 2021, Morcillo et al. 2016) including design rules¹³. We examined some production factors like video length, genre, format and audio - said to have a great impact on the popularity and added some more production and general factors. For example, was the corresponding paper named in the video and linked permanently with a DOI and was the research data visualized in the video?

Formal metrics for each video:

- a) Subjects
- b) License information (Licence Type e.g. CC BY, CC BY ND)
- c) Title / Intro (yes/no)
- d) Credit Paper (yes/no)
- e) Credit Research Data (yes/no)
- f) DOI (yes/no)

Metrics for Design and Production:

- g) Visualized research data included (yes/no)
- h) Length (1-4 minutes, 4-5 minutes, >5 minutes)
- i) Genre (Documentary Style, Animation, Dynamic Presentation, Monologue/Interview, Simple presentation)
- j) Additional Elements (Text, Graphics, Animation, Presenters, Interviews, experiment / lab footage, documentary / real life footage)
- k) Sound Quality (good, medium, bad, no sound)

¹² Data see: <https://doi.org/10.17026/dans-xh6-fama>

¹³ <https://iopscience.iop.org/journal/1538-3873/page/video-abstract-guidelines>

Results

Formal Criteria

Subjects

Our subset included nineteen videos from Physics, nine from Earth Science and one each from Informatics, Biology, Information Science, Engineering, Environmental Science.

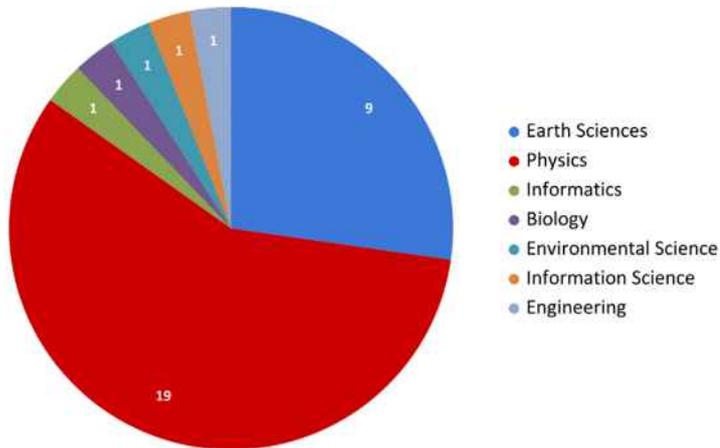


Fig. 1: Subjects (Sample:33)

Licenses

All videos have clear licence information. Nineteen videos were published under CC BY 3.0 DE, twelve under CC BY 3.0 unported, one by CC NC, one CC ND and one CC NC/ND.

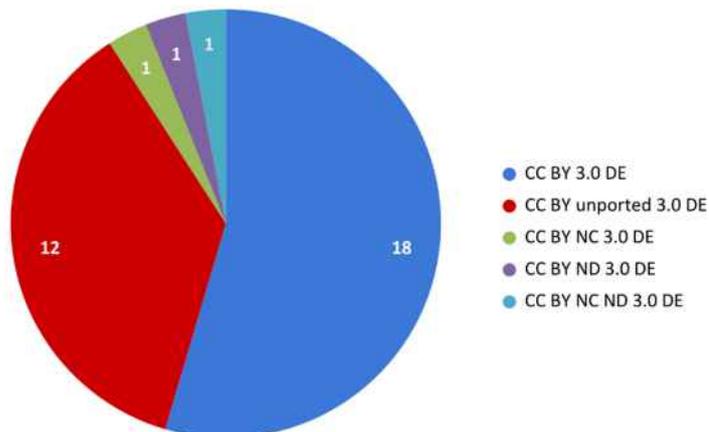


Fig. 2: Licenses (Sample:33)

Intro or Title

Thirty video abstracts had an intro or at least a clear title, three did not.

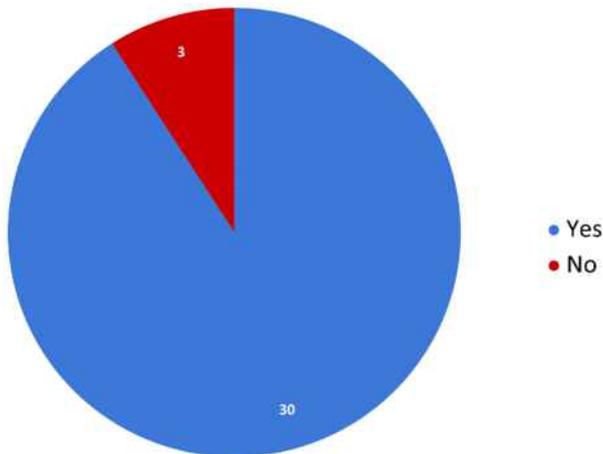


Fig. 3: Intro or Title (Sample:33)

Credit / Paper / Research Data

Only eighteen videos included credits for the paper (either in form of the title or DOI), ten did not have any information or a link to the corresponding paper or article. None of the VA added a link to the underlying research data.

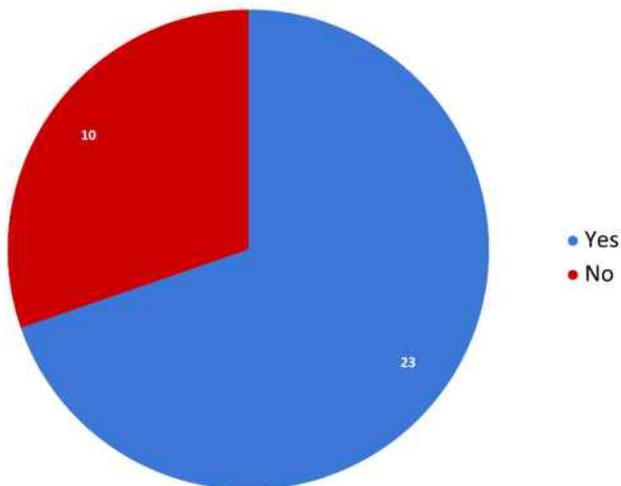


Fig. 4: Credit Paper (Sample:33)

Design Features and Production

Have the authors used scientific data visualizations?

Twenty-five video abstracts included a visualization of their scientific data in the form of models, graphs etc. eight did not make their data visible in their video abstracts.

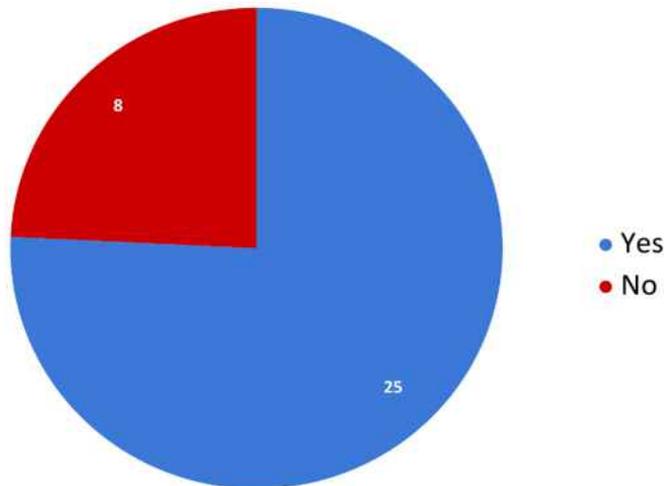


Fig. 5: Scientific Data (Sample:33)

Length

The average length of the video abstracts was 4 minutes.

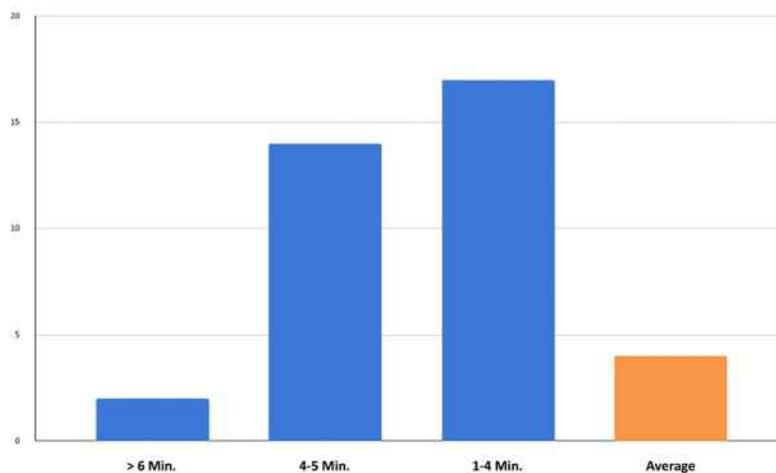


Fig. 6: Length and Average (Sample:33)

Format / Genre

Sixteen VAs used a simple presentation (scenes featuring the authors of the scientific article, category includes all kinds of author(s)' appearance on screen, author(s) talking to the camera, interviews)

Eight used a documentary style (scenes recorded with normal camera, showing the subject of the research acting, footage, reconstruction of experiments, footage demonstrating lab and/or fieldwork)

Four used a dynamic presentation. As the name suggests, the focus is on a dynamic style. The authors used several different materials (animation, experiment/observation, documentary/real life, etc.) and made greater use of driving background music. Two videos of that type are under 2:30 min long.

Three used animations (scenes generated with non-linear or analog animation techniques, category includes whiteboard-, stop motion-, 3D-, timelapse animations, as well as screen recordings).

One video abstract was designed in the form of a conversation or interview. The author is interviewed and tells about her research work. In between, scenes of the computer monitor with the open document, the paper, can be seen. The computer monitor was filmed by the camera.

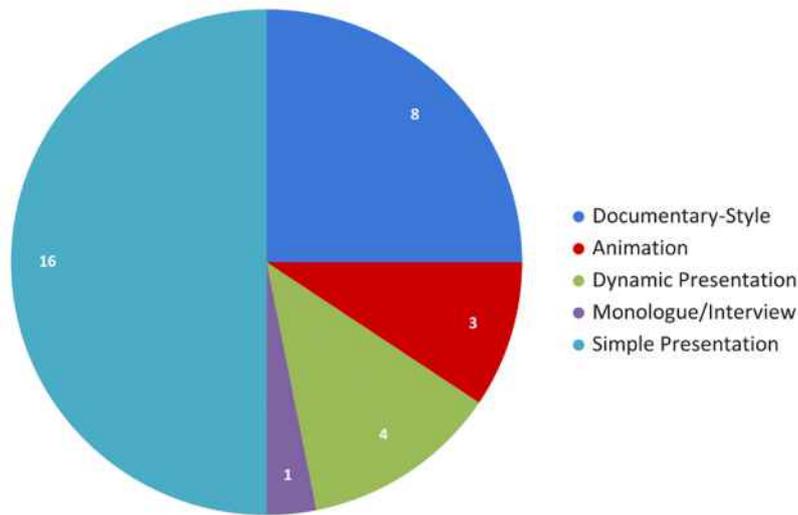


Fig. 7: Format/Genre (Sample:33)

Additional Elements

Additional elements include text (25), graphics / still Images (27), animations (11), scenes featuring authors / presenters (16), interviews / talking heads (16), scenes featuring experiments (8) and documentaries / real life scenes (12).

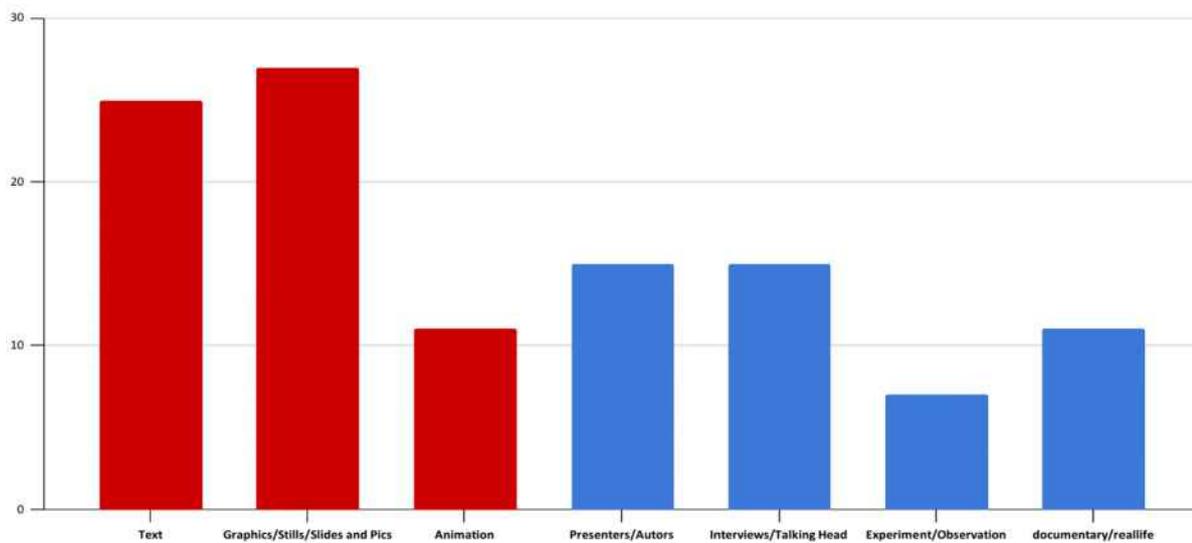


Fig. 8: Additional Elements (Sample:33)

Audio-Quality

Seventeen video abstracts had a good audio quality for the verbal narration, ten a medium, three a bad. Three VA did only use music and not a narration.

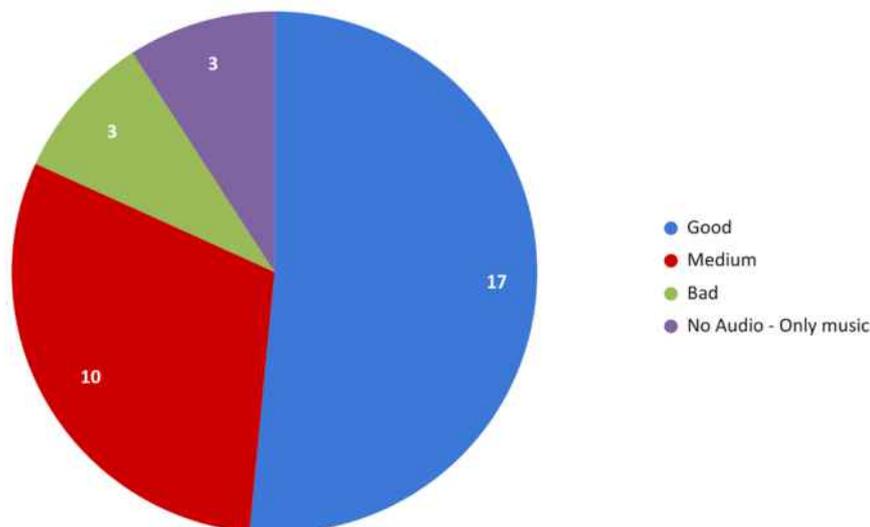


Fig. 9: Audio-Quality (Sample:33)

Conclusion

Since all VAs were published in the science compliant TIB AV portal, all had a DOI, license information, author information, title, publication year, and clear subject attribution in form of metadata.

On average, the videos from our analysis were not longer than 4 minutes (3:55 min.). Most VA did include additional relevant material such as images, animations, and lab or real footage. The majority of the VA entries had medium to good sound quality (82%) and were thus acoustically acceptable. Three videos had no title visible in the video. Most video abstracts used just one presenter/speaker who gave a “classic” simple short-presentation. In these VAs the author is shown on the screen, talking to the camera or giving an interview. Documentary style VAs were the second most common style. These types added animation, experiment/observation, documentary/real life footage and/or background music. Only half of the authors linked to their papers and only three quarters included a visualization of their research data. None of them linked to the underlying research data.

In principle the three major points, which occur in most of the guidelines are followed by the authors. These are:

- A video abstract should not last longer than 4 minutes.
- Inclusion of additional relevant material such as images, animations, and lab footage is strongly encouraged.
- A video abstract must include a soundtrack providing a clear verbal narration of the visual content.

However the results of the study show that there is still a lot of potential in improving video abstracts, especially with regard to formal criteria (e.g. visible title) as well as linking with the according papers and the underlying research data. Strictly speaking video abstracts that do not have a recognizable reference to a scientific article (Credit Paper) cannot be assigned to this genre. Here it would have to be checked whether the video abstract is related to a paper at all.

This data, combined with future quantitative and qualitative research, will hopefully provide new insights into the global study of audio-visual communication of science and help the authors to get more out of their video abstracts. Since we have seen many videos that still have a lot of potential for improvement, we would like to give the following advice to the authors and hope that in the future these will also be taken into account in the guidelines.

- Make sure your VA has a DOI, so that it is sustainably citable
- Make sure your VA has information on the licence for reuse (e.g. Creative Commons¹⁴)
- Make sure your VA includes credits for your paper, best link it via DOI
- Make sure you link to the underlying research data
- Make sure your VA has an Intro and an Outro,
- Make sure you mention all of the authors
- Make sure you share your video on at least one science compliant portal (e.g. TIB AV-Portal), so that it is citable, digitally preserved and connected to research data and paper.

Sample (Research Data)

Full list of video abstracts

Nr.	Title	DOI (Video)
1	Soil lacquer peel DIY: simply capturing beauty	10.5446/40805
2	A system's wave function is uniquely determined by its underlying physical state	10.5446/21893
3	Overview on the open research knowledge graph	10.5446/52261
4	Dynamics of salt intrusion in the Mekong Delta	<u>10.5446/53547</u>
5	TROPOMI measurements and WRF CO modelling to understand extreme air pollution events in India	<u>10.5446/50921</u>
6	Communicating climate change in a 'post-factual' society: Lessons learned from the Pole to Paris campaign	10.5281/zenodo.2662144
7	SPHY-MMF, Coupled Hydrology-Soil Erosion Model	<u>10.5446/37585</u>
8	Transition from electromagnetically induced transparency to Autler–Townes splitting in cold cesium atoms	<u>10.5446/38882</u>
9	A "mental models" approach to the communication of subsurface hydrology and hazards	<u>10.5446/21332</u>
10	Scientific Audiovisual Materials and Linked Open Data	<u>10.5446/17789</u>
11	Magnetically guided multi-coordinate positioning system	<u>10.5446/39650</u>

¹⁴ <https://creativecommons.org/>

12	Exciton effective mass enhancement in coupled quantum wells in electric and magnetic fields	<u>10.5446/38806</u>
13	Coulomb blockade model of permeation and selectivity in biological ion channels	<u>10.5446/38744</u>
14	Efficiency of the SQUID ratchet driven by external current	<u>10.5446/38751</u>
15	Anderson localization of composite excitations in disordered optomechanical arrays	<u>10.5446/21891</u>
16	Groundwater fluctuations during a debris flow event in Western Norway – triggered by rain and snowmelt	<u>10.5446/53450</u>
17	Spatiotemporal Variability in the Oxidative Potential of Ambient Fine Particulate Matter in Midwestern United States	<u>10.5446/54718</u>
18	Noncommutative correction to the entropy of BTZ black hole based on Lorentzian mass distribution with GUP	<u>10.5446/49798</u>
19	An integrated observation dataset of the hydrological-thermal deformation in permafrost slopes and engineering infrastructure in the Qinghai-Tibet Engineering Corridor	<u>10.5446/53800</u>
20	First Principle of the electronic and optical properties of transition metal dicalcogenide (tmd)	<u>10.5446/49800</u>
21	Feed conversion efficiency in aquaculture: do we measure it correctly?	<u>10.5446/39356</u>
22	On the relative intensity of Poisson's spot	<u>10.5446/38422</u>
23	Voice-Sensitive Regions in the Dog and Human Brain Are Revealed by Comparative fMRI	<u>10.5446/32194</u>
24	Cyclotron and Synchrotron: some applications	<u>10.5446/50119</u>
25	Exploring dispersal barriers using landscape genetic resistance modelling in scarlet macaws of the Peruvian Amazon	<u>10.5446/32196</u>
26	EPR-based ghost imaging using a single-photon-sensitive camera	<u>10.5446/39018</u>
27	The role of edible mushrooms in the green synthesis of CdS quantum dots	<u>10.5446/49858</u>
28	Obtaining the Feynman path integral through the brownian motion description	<u>10.5446/49804</u>
29	Conceptual design of the beam source for the DEMO Neutral Beam Injectors	<u>10.5446/21890</u>
30	Assessing ExxonMobil's climate change communications (1977–2014)	<u>10.5446/39391</u>

31	The search for Bose–Einstein condensation of excitons in Cu ₂ O: exciton-Auger recombination versus biexciton formation	10.5446/38742
32	Structural analysis and phase transition study of the compound Pb(x) Cd(1-x)TiO ₃ : by X-ray diffraction and Raman spectroscopy	<u>10.5446/50261</u>
33	Development of the RFI monitoring application for the BINGO radio telescopic site	<u>10.5446/50109</u>

Literature

- Berkowitz, J. (2013). Video abstracts, the latest trend in scientific publishing: Will “publish or perish” soon include “video or vanish”? University Affairs. Retrieved from <http://www.universityaffairs.ca/features/feature-article/video-abstracts-the-latest-trend-in-scientific-publishing/>
- Boy, B., Bucher, H.-J., Christ, K. (2020). Audiovisual Science Communication on TV and YouTube. How Recipients Understand and Evaluate Science Videos. IN *Frontiers in Communication*. Vol. 5 <https://doi.org/10.3389/fcomm.2020.608620>
- Bucher, H.-J. (2020). The contribution of media studies to the understanding of science communication, in *Handbook of Science Communication*, eds A. Leßmöllmann, M. Dascal, and T. Gloning (Boston, MA; Berlin: De Gruyter), 51–76.
- Ferreira M. et al. (2021). Audio-Visual Tools in Science Communication: The Video Abstract in Ecology and Environmental Sciences. IN *Frontiers in Communication*. February 2021. DOI: 10.3389/fcomm.2021.5962
- Garcia-Aviles, J.A., de Lara, A. (2020). “An Overview of Science Online Videos”. In *Communicating Science and Technology through Online Video. Researching a New Media Phenomenon*, eds B. León, and M. Bourk (New York, NY; London: Routledge)
- León, B., and Bourk, M. (2020). “Investigating science-related online video.” In *Communicating Science and Technology through Online Video. Researching a New Media Phenomenon*, eds B. León, and M. Bourk (New York, NY; London: Routledge)
- Molnár, A.D. (2016). How to turn an abstract into a video abstract, TPD 2016 - Workshop: Videos in digital libraries: What’s in it for libraries, publishers and scientists? 2016. <https://doi.org/10.5446/19853>
- Morcillo, J. M., Czurda, K., and Trotha, C. Y. R. V. (2016). Typologies of the popular science web video. *J. Sci. Commun.* 15 (4), 32. doi:10.22323/2.15040202
- Plank, M., Marín-Arraiza, P., Molnár D.A.(2018). Video Abstract for Science Education. In: *Exploring the Micro, Meso and Macro – EDEN Annual Conference Proceedings, 2018, Genova*, (S.123-130)
- Plank, M., Molnár, A., Marín Arraiza, P., (2017). Extending Media Literacy Education. Proceedings of IFLA WLIC 2017 <http://library.ifla.org/1776/1/242-plank-en.pdf>
- Spicer, S. (2014). Exploring Video Abstracts in Science Journals: An Overview and CaseStudy. *Journal of Librarianship and Scholarly Communication*, 2(2), p. eP1110. doi: <http://doi.org/10.7710/2162-3309.1110>
- van Edig, X. (2016). Video abstracts and video supplements to scientific articles – experiences from Copernicus Publications. doi: <https://doi.org/10.5281/zenodo.59819>
- van Norden, R. (2015). Interdisciplinary research by the numbers. *Nature*, 525(7569), 306-307. Retrieved from <http://www.nature.com/news/interdisciplinary-research-by-the-numbers-1.18349>
- Whitesides, G. (2011, April 29). Publishing Your Research 101. Impact of video on scientific articles. American Chemical Society. [Video]. Retrieved from: <https://www.youtube.com/watch?v=HboNzrq0MKE>
- Zong, Q., Xie, Y., Tuo, R. et al. The impact of video abstract on citation counts: evidence from a retrospective cohort study of *New Journal of Physics*. *Scientometrics* 119, 1715–1727 (2019). <https://doi.org/10.1007/s11192-019-03108-w>

Grey Literature and Persistent Identifiers: GreyNet's Use Case

Dominic Farace, GreyNet International

Stefania Biagioni and **Carlo Carlesi**, GreyGuide ISTI-CNR, Italy

Chris Baars, DANS-KNAW, Netherlands

Abstract

The PID (Persistent Identifier) Project is a follow-up to the AccessGrey Project¹ carried out in 2019 in which an online survey was held among stakeholders in GreyNet's community of practice. Recipients were asked their opinions about persistent identifiers and grey literature. The focus now in this project is expanded to include the DOI for research outputs alongside the ORCID for authors/researchers, and the ROR ID for research organizations. This project seeks to go beyond a straightforward compilation and linking of these PIDs by building the PID Graph and contribute to other PID-Graphs built by service providers like OpenAIRE². In this case, the PID Graph seeks to demonstrate how persistent identifiers can further research in the field of grey literature; and, how they contribute in making research entities conform to the FAIR data principles: Findable, Accessible, Interoperable, and Reusable. PIDs and the PID Graph are also seen to serve in the digital transformation of grey literature and as such will contribute to education and training in this field of information. DataCite Commons³ used in this project is a web search interface for the PID Graph. The results from queries directed to the PID Graph produced in this project will not only serve as a use case for GreyNet but will also provide a model for other communities of practice in grey literature.

Chapter Outline

1. Background – AccessGrey Project and Persistent Identifiers
2. Components and Data Workflow in GreyNet's PID Project
3. Implementation of the PID Graph
4. Persistent Identifiers and the PID Graph conform to FAIR data principles
5. Some Conclusions drawn from the PID Project

1. Background – AccessGrey Project and Persistent Identifiers⁴

In 2019, an online survey was carried out among GreyNet's community of practice in order to gain their opinions on the uses and applications of persistent identifiers for grey literature. Results from an online survey within the AccessGrey Project clearly indicate a positive opinion about persistent identifiers for grey literature. Emphasis in the AccessGrey Project focused on two of the four collections in the GreyGuide Repository – the GLP Collection of conference papers and the new RGL Collection of multiple grey literature document types. The search of records in the GLP collection enabled formulation of the questions used in the online survey pertaining to persistent identifiers in particular the DOI.

1.1 AccessGrey Questionnaire⁵ - 'Persistent Identifiers and Grey Literature'

- Q1. Persistent identifiers increase access to grey literature
- Q2. Persistent identifiers serve as an incentive in the acquisition of grey literature
- Q3. Persistent identifiers increase the citation of grey literature
- Q4. Persistent identifiers allow for the preservation of grey literature
- Q5. Persistent identifiers are vital in linking and cross-linking data
- Q6. A DOI is a quality indicator that increases the value of grey literature

- Q7. A repository or data archive that assigns DOIs to metadata records is more likely to attract content providers
- Q8. Do you have an ORCID or another author/researcher unique persistent identifier?
- Q9. Does one or more of your publications have an assigned DOI?

Survey Population: 509		Survey Respondents: 56		Survey Results: 11%	
	Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree
Q1	31	19	6	0	0
Q2	17	22	15	2	0
Q3	31	19	5	6	1
Q4	23	23	8	2	0
Q5	33	15	7	1	0
Q6	17	19	13	6	1
Q7	19	26	7	3	1
	Yes	No	Non-Applicable		
Q8	37	13	6		
Q9	40	9	5		

Table 1: AccessGrey Survey Results

The results of the questionnaire, which constituted the first part of the project, were significantly positive regarding persistent identifiers and grey literature.

1.2. DOI an Incentive for Acquisitions

The results from the second part of the AccessGrey project however did not indicate that the minting of DOIs for research outputs would be a sufficient incentive for their acquisition in a repository, namely one that relies on self-archiving. During that project, new records entered in the RGL (Resources in Grey Literature) collection received a DOI and a system generated citation. However, fewer full-text metadata records were harvested during this part of the project than expected. This perhaps coincides with the response to the survey question (Q2) in which nearly 27% of the respondents were uncertain whether persistent identifiers serve as an incentive in the acquisition of grey literature.

2. Components and Data Workflow in GreyNet’s PID Project

A persistent identifier (PID) is a permanent reference and unique label to an object that is independent of the storage location. The unique label ensures that the object can always be found, even if the name of the object or the repository changes. As a result, an object can always be found unambiguously on the basis of its PID. This is important for the long-term storage (archiving) of objects in a rapidly changing world.⁶

In short, PIDs

1. Provide the address of an object such as a landing page in a repository
2. Can be used to link objects and in so doing connect other associated metadata in a record
3. Unambiguously Identify objects even if they move to other systems and services
4. And are computer readable, demonstrating their interconnectedness with other research communities

The value of the PID (persistent identifier) not only provides a link to a digital object but also allows the metadata associated with the digital object to become connected. When that metadata itself is expressed as a PID, this further allows for the creation of a PID Graph that models FAIR data principles: Findable, Accessible, Interoperable, and Reusable. These four principles will be discussed further in the chapter, following the introduction of GreyNet's network service in which they are applied.

2.1 GreyNet International, Grey Literature Network Service⁷

In order to obtain optimal use of persistent identifiers a sustained data infrastructure must be in place within a community of practice, one that facilitates a coherent data workflow. It is only then that a PID Graph can be constructed and implemented. In this section, we look at the various components in GreyNet's data infrastructure and then discuss how they are implemented within its workflow. It is important to mention here that GreyNet's workflow as it applies to this project includes retrospective input.

2.2 Components of the Data Infrastructure integrated in the PID Project

■ GreyGuide Repository⁸ and Portal to Good Practices and Resources in Grey Literature

GreyNet International collaborated with ISTI-CNR to construct the GreyGuide a web access repository, which would come to house its collections of accepted conference proposals (GLA), published conference papers (GLP) and author-researcher biographical records (BIO). To this end, open-source software was identified and incorporated, metadata templates were created to fit the three document types, and in 2017 these collections were fully online accessible.

■ OpenDOAR, Directory of Open Access Repositories⁹

The GreyGuide Repository is registered in the OpenDOAR Directory of Open Access Repositories¹⁰. It can be mentioned here that while the GLA and BIO collections in the GreyGuide Repository rely on self-archiving, records in the GLP collection are entered by the system manager.

■ DOI, Digital object identifier¹¹ and DataCite.org¹²

In 2018, GreyNet became a DOI minting service within DataCite and began assigning DOIs to its collection of Conference Papers in the GreyGuide Repository. Since it is this collection upon which our PID Project is based, the ORCID and ROR IDs had to be included in the DOI records in order later construct the PID Graph and be part of other PID Graphs (e.g., DataCite Commons, OpenAIRE).

■ ORCID, Open Researcher and Contributor ID¹³

Also, in that same year, ORCIDs were included in biographical records in the GreyGuide; and, an active campaign among GreyNet's authors and researchers was initiated – encouraging them to register an ORCID if they did not yet have one. In order to facilitate this, a link to the ORCID registry was provided¹⁴.

■ ROR, Research Organization Registry ID¹⁵

In 2020 the ROR ID for research organizations was added as a metadata field in BIO records in the GreyGuide. By way of a search in the ROR Registry, ROR IDs of organizations could be online accessed and included in the records of those authors and researchers, whose conference papers are archived in the GLP collection as well as in their corresponding DOI records in DataCite. GreyNet has since applied for a ROR ID and awaits its assignment. A ROR ID unlike an ORCID is not assigned separately but rather in interval batch-releases, once new records have been approved.

It is worthwhile to note that the ROR ID of an organization linked with a research output such as a conference paper or other grey literature document type might be perceived as

a quality indicator. If we look back to question (Q6) in the AccessGrey Survey, over 23% of the respondents were uncertain whether the DOI is a quality indicator that increases the value of grey literature. However, if DOIs were connected to their corresponding ORCID and ROR IDs, there might be less uncertainty.

2.3 The Data Workflow implemented in the PID Project

Our PID Project team was formed bringing together human resources and expertise needed, namely the system management and development of the GreyGuide Repository; the communication and network management of the GreyNet Community, and the acquired knowledge and experience of the PID Graph. From early January 2021 through the first week of March 2021, GreyNet undertook three tasks integral to the PID Project.

First, to complete minting DOIs for its collection of conference papers in the GL-Series including those published in 2021. The collection now totals 443 conference papers with DOIs in DataCite that accounts for the population of our project. Other service providers, namely DANS EASY¹⁶ for GreyNet’s published datasets and the TIB AV Portal¹⁷ for its conference video presentations also have assigned DOIs in DataCite; however, these are not included in the population of the project.

The **second** task that ran parallel with the first was the retrospective search and retrieval of ORCID and ROR IDs that were added to both the DOI metadata records and their respective BIO records in the GreyGuide Repository. The retrospective task also included the input of biographical records on behalf of authors and researchers whose names appear in the GLP collection, but who had not yet submitted a BIO record. This was accomplished in part by retrieving biographical notes from previous conferences in the GL-Series preserved in GreyNet’s inhouse archive and partly via Google searches.

A **third** ongoing task dealt with records that needed some modification in order to benefit the PID Project, such as

- (1) An existing ORCID in a record carries 16 digits but is not proceeded by [https://orcid.org/] and as such is not actionable; (2) An ORCID is retrieved only to find the message ‘No public information available’, which makes it difficult if not impossible to confirm the identity of the author/researcher; and (3) When an author or researcher’s organization is absent or unclear in a record, it becomes difficult or is not possible to assign a ROR ID using the ROR Registry.

While these and other such problems were few in number, the time required to correct them was disproportionate. Nevertheless, when a system and service rely on self-archiving and when a persistent identifier such as the ORCID can only be acquired by the author-researcher – him or herself, then these tasks must be calculated in the workflow.

2.4 Compilation of Actionable Persistent Identifiers

Now that the complete collection of conference papers in the GL-Series has an assigned a DOI in DataCite, which incorporates their corresponding ORCID and ROR IDs, the number of actionable persistent identifiers for our project is accounted for. And, this then allows for the construction of the PID Graph.

Conference Papers	Authors-Researchers	Research Organizations
GLP Collection: 443	BIO Collection: 238	BIO Collection: 238
 443	 146	 180
100%	61.3%	75.6%

Table 2: Actionable PIDs compiled in the Project (as of *March 13, 2021*)

3. Implementation of the PID Graph

In an article published in early January 2021, GreyNet's attention was drawn to the benefits of connecting the various types of persistent identifiers in producing a PID Graph¹⁸. For our project, this includes the DOI, ORCID, and ROR ID. It is expected that this PID infrastructure would further demonstrate the value of persistent identifiers and open the potential for more research - in our case, research in the field of grey literature.

To construct the PID Graph two elements are required:

- (1) backend services that collect PID connections in a standardized way focusing on two PIDs that are connected. This is essentially building the elements of the graph;
- (2) query interfaces that combine these connections with PID metadata. A technology that is highly suitable is GraphQL¹⁹. GraphQL is an open-source data query and manipulation language for APIs, and a runtime for fulfilling queries with existing data. This widely adopted query language provides a standardized interface that can be federated, making it easier to build client applications for the PID Graph. Applications built on top of the PID Graph allow users to explore the rich connections between PIDs and to address specific use cases. The PID Graph demonstrates that we can gain more from PIDs when we look at their connections – indicating that the sum is more than its parts.

3.1 Examples of the PID Graph

Below are examples of four PID Graphs drawn from GreyNet's store of persistent identifiers. Each graph is comprised of multiple resources (nodes) that are connected by lines (edges). In the first diagram, the PID Graph appears in horizontal format and depicts from a DOI perspective three publications connected with the authors and their respective organizations. In the second diagram, the PID Graph appears in cluster format and depicts from a DOI perspective the same three publications as in the first diagram; however, now they are connected with the authors' names and their respective organizations. In the third diagram, the PID Graph depicts from an ORCID perspective an author and his respective organization linked to seven publications. One of the publications is further linked to three co-authors of whom only one organization is shown. And, in the fourth diagram, from a ROR ID perspective – a research organization is encircled and linked to a cluster of publications that is further encircled and linked to a number of authors. Three of the authors appear also linked to their own respective organization.

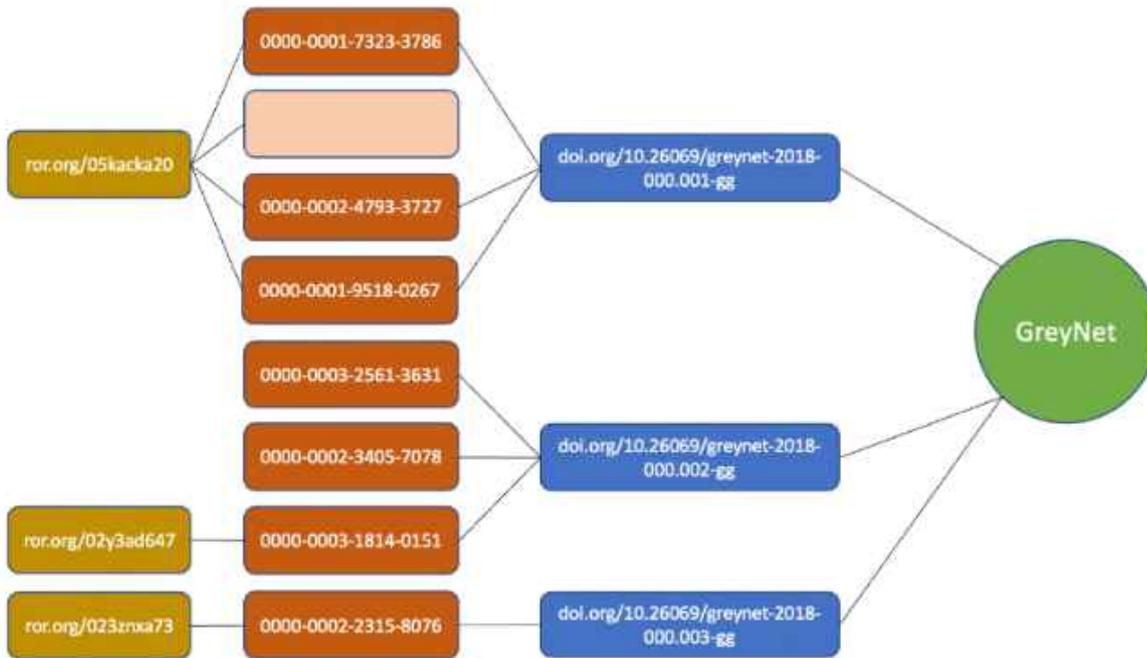


Diagram 1: PID Graph from a DOI perspective in horizontal format

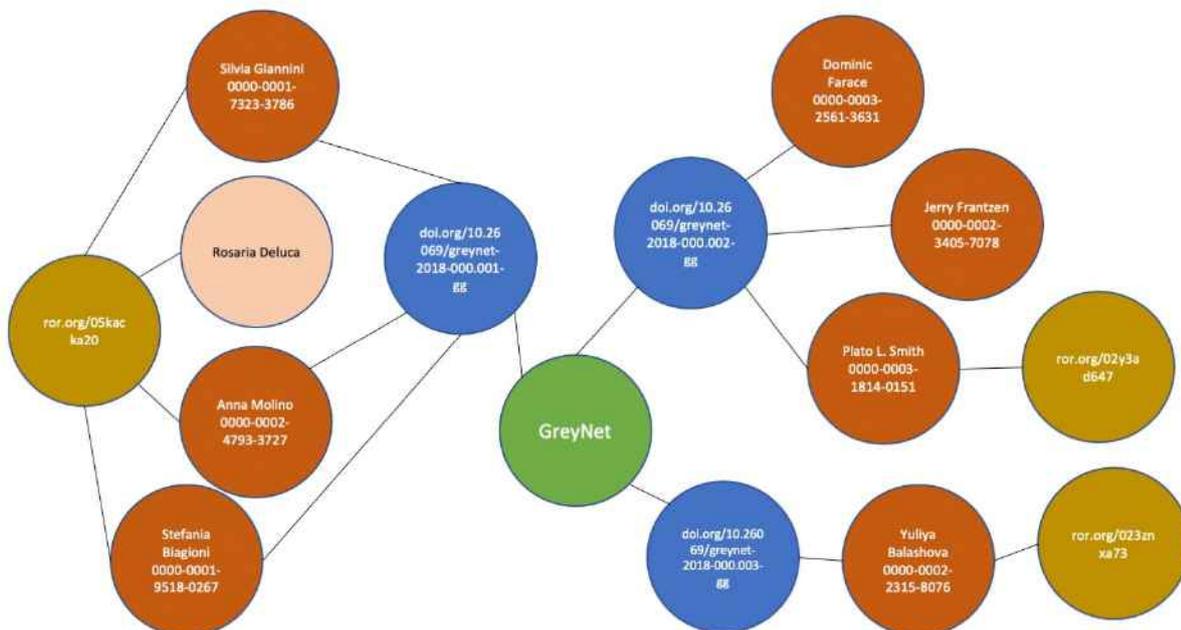


Diagram 2: PID Graph from a DOI perspective in cluster format

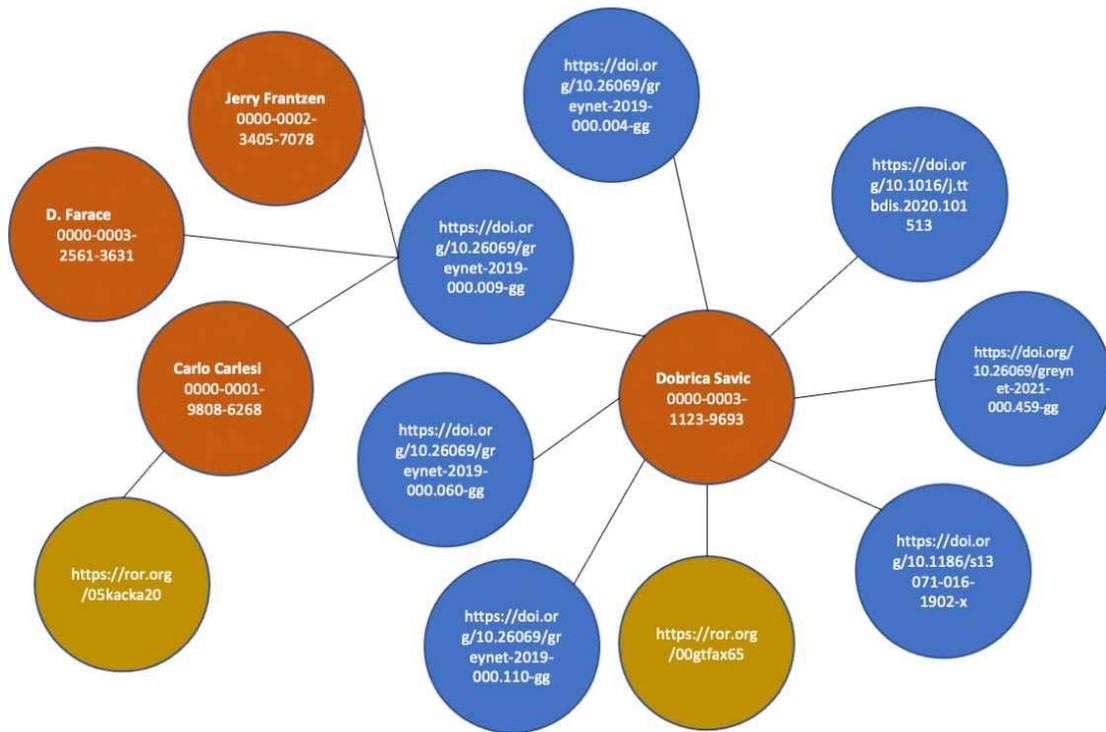


Diagram 3: PID Graph from an ORCID perspective in cluster format

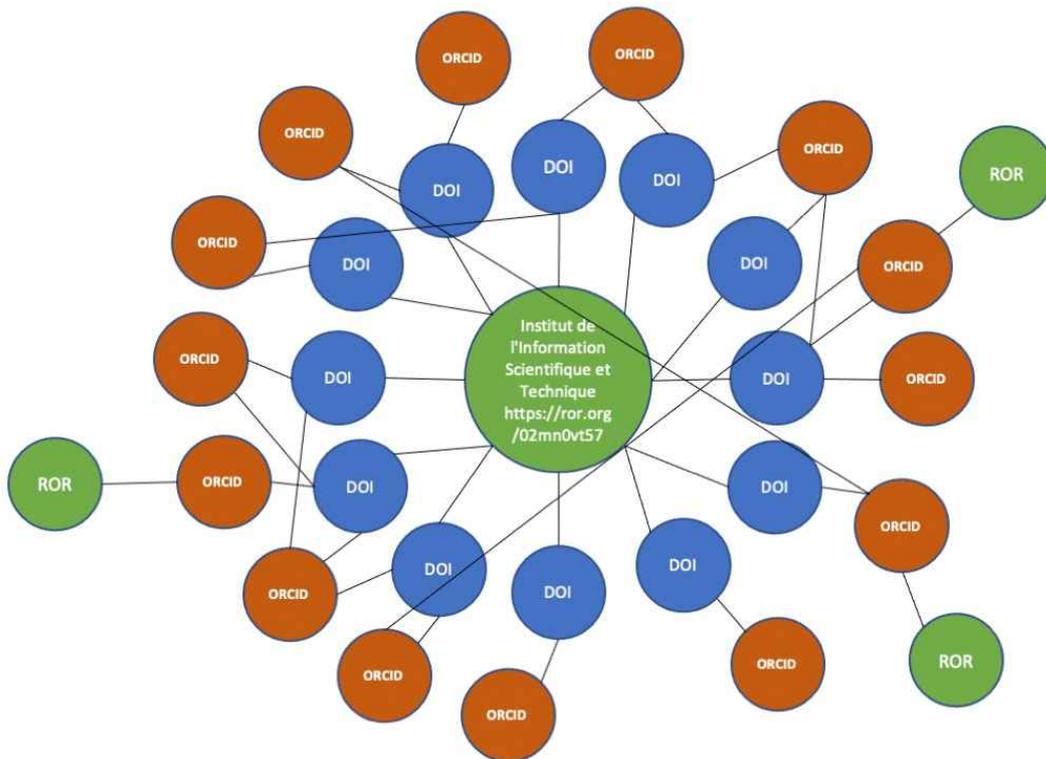


Diagram 4: PID Graph from a ROR ID perspective in cluster format

In the PID Graph, persistent identifiers are themselves the basic entities that are linked together; whatever they refer to is left implicit. This approach requires that the PID metadata are sufficiently rich to represent the relationships of interest and that the PIDs are of high enough quality. The advantage is that it becomes much easier to create graphs and to implement and scale rather than working with concepts and knowledge extraction.

4. PIDs and the PID Graph in relation to FAIR data principles

PIDs themselves allow for the guarantee of interconnected services from minting to linking onto access and preservation. When these services are situated in the workflow of a mature community of practice, they create a FAIR research environment.

[An extract abridged and revised from ‘Connected Research: The Potential of the PID Graph’]²⁰

PIDs contribute in making research entities conform to the FAIR data principles; Findable, Accessible, Interoperable, and Reusable. By way of the PID Graph connections between different entities within the research landscape allow researchers to access new information. PIDs also play a role in the reusability of data by enabling rich metadata and their provenance to be associated with a digital object. PIDs provide the possibility to link entities long-term and enable information exchange by identifying persons and organizations over different services.

The overall PID infrastructure is made up of PID service providers, repositories, curation systems, aggregators, indexes, metadata, standards, and people. PIDs connect all of these elements, not only technically, via metadata and integrations, but also socially, via communities that have formed over decades or longer. The table below identifies the various types of PIDs and the maturity of their infrastructure. Since 2018, the ROR ID has moved from an emerging entity to a mature one.

Research Entity	PID Types Used	Maturity of PID Infrastructure
Publication	DOI, accession number, handle, URN, Scopus EID, Web of Science UID, PMID, PMC, arXiv identifier, BibCode, ISSN, ISBN, PURL	mature
Citation	OCI (secondary aggregation of information)	emerging
Conference	DOI, accession number	emerging
Researcher (or scholar)	ORCID IDs, ISNI (also DAIs, VIAFs, arXivIDs, OpenIDs, ResearcherIDs, ScopusIDs)	mature
Organization	DOI, ISNI, GRID, Ringgold IDs, ROR IDs	emerging
Data	DOI, accession number, handle, PURL, URN, ARK	mature

Table 3: PID Types and Infrastructure Maturity

GreyNet International now in its 28th year can be considered a mature research community socially. By including PIDs for objects, projects, persons, and organizations in the metadata, the technical maturity of GreyNet’s infrastructure can now likewise be demonstrated. As a result of this ① sustainable connections can be made ② objects, projects, persons, and organizations become computer readable and understandable by other services like DataCite and OpenAIRE ③ A PID-Graph can be created and GreyNet information can also become part of other PID-graphs ④ Other services, like OpenAIRE PID-Graph and DataCite Commons²¹ can be used to query or for purposes of analysis, and ⑤ It is a demonstration of FAIR-principles for grey literature. To include and expand on the FAIR principles, PIDs and metadata help ensure that the entities they refer to are

- usable and citable: pointing directly to an object, such as a specific item or a specific version of a dataset; hence increasing the usability of that object for researchers. It also helps them formally cite research outputs such as data and resources, which in turn facilitates reuse and helps increase recognition.
- Assessable: PIDs enable reliable measurement and prediction of impact, facilitating a more strategic approach to investment, driving maximum benefit, and ensuring that valuable resources are sustained.

5. Some Conclusions Drawn from the PID Project

Research in the field of grey literature will likely increase due to the incorporation and use of persistent identifiers. PIDs like other rich metadata can be counted and cross tabulated, enabling researchers to examine relationships in and among diverse types of data. As such, PIDs are actionable and can be used for new research. Furthermore, PIDs and the PID Graph can be seen not only to serve research in grey literature but also extend to new services in areas of education and training.

PIDs and the PID Graph are shown to have real value in defining GreyNet's position as a mature research organization by sustaining and leveraging its resources, by adhering to the FAIR data principles, and by signaling increased trust in grey literature beyond its own community of practice.

While the minting of a DOI was not of itself a sufficient selling point in the earlier AccessGrey Project for attracting content to a repository, the DOI now linked to the ORCID and ROR IDs illustrated by the PID Graph may prove more effective. Also, while the AccessGrey Project laid the groundwork and direction for this PID Project, it is our understanding that implementation of the PID Graph will go even further to provide a new strategy and approach to research in the field of grey literature.

Linked References

- ¹ <https://doi.org/10.17026/dans-zzf-cje3>
- ² <https://graph.openaire.eu/>
- ³ <https://commons.datacite.org/>
- ⁴ <https://doi.org/10.26069/greynet-2020-000.219-gg>
- ⁵ <https://doi.org/10.17026/dans-zzf-cje3>
- ⁶ https://nl.m.wikipedia.org/wiki/Persistent_identifier
- ⁷ <http://www.greynet.org/>
- ⁸ <http://greyguiderep.isti.cnr.it/>
- ⁹ <https://v2.sherpa.ac.uk/opensoar/>
- ¹⁰ <https://v2.sherpa.ac.uk/id/repository/9690>
- ¹¹ https://en.wikipedia.org/wiki/Digital_object_identifier
- ¹² <https://datacite.org/value.html>
- ¹³ <https://en.wikipedia.org/wiki/ORCID>
- ¹⁴ <https://orcid.org/register>
- ¹⁵ <https://ror.org/>
- ¹⁶ <https://easy.dans.knaw.nl/ui/home>
- ¹⁷ <https://av.tib.eu/>
- ¹⁸ <https://doi.org/10.1016/j.patter.2020.100180>
- ¹⁹ <https://graphql.org>
- ²⁰ [https://www.cell.com/patterns/pdf/S2666-3899\(20\)30244-0.pdf](https://www.cell.com/patterns/pdf/S2666-3899(20)30244-0.pdf)
- ²¹ <https://commons.datacite.org/>

Selected Resources

Introducing the PID Graph

<https://www.project-freya.eu/en/blogs/blogs/the-pid-graph>

The FREYA project

<https://www.project-freya.eu/en>

Connected Research: The Potential of the PID Graph

[https://www.cell.com/patterns/pdf/S2666-3899\(20\)30244-0.pdf](https://www.cell.com/patterns/pdf/S2666-3899(20)30244-0.pdf)

The power of PIDs: Using persistent identifiers to link research outputs in the Netherlands

<https://www.dpconline.org/blog/the-power-of-pids>

NARCIS enriched with the first results of a PID-Graph

<https://dans.knaw.nl/en/current/news/narcis-enriched-with-the-first-results-of-the-pid-graph>

AccessGrey: Securing Open Access to Grey Literature for Science and Society

<http://greyguiderep.isti.cnr.it/linkdoc.php?idcode=2020-GL21-007&authority=GreyGuide&collection=GLP&&langver=en>

AccessGrey Online Questionnaire

[file:///C:/Users/GreyNet/Downloads/Survey%20Results%20\(Anonymous\)%20\(4\).pdf](file:///C:/Users/GreyNet/Downloads/Survey%20Results%20(Anonymous)%20(4).pdf)

Data from "AccessGrey: Securing Open Access to Grey Literature for Science and Society"

<http://greyguiderep.isti.cnr.it/linkdoc.php?idcode=2020-RGL01-002&authority=GreyGuide&collection=RGL&&langver=en>

The ISTI Library was founded in 1954. Since 2000 is part of the Library Service and Scientific Documentation Centre of the CNR Research Area in Pisa.

The Library works to support the research activities and the researchers. The staff takes care of and enhances the bibliographic and documentary heritage of ISTI and the other institutes belonging to the CNR Research Area.

The Library is involved in the fields of Grey Literature (GL), Digital Libraries, and institutional archives. The Library subscribes and promotes the principles and actions of the Open Science community, giving information about national and international initiatives and managing services for the institutes located in the Area and the interested authors.

The Library staff manages the ISTI scientific production and supports the authors in the research assessment procedures.

The descriptions of ISTI laboratories research activities are extracted from ISTI laboratories brochures. Other info are available here.

Istituto di Scienza e
Tecnologie dell'Informazione
"A. Faedo" (ISTI)
Via G. Moruzzi 1 - 56124
Pisa - ITALY
Contacts:
library@area.pi.cnr.it

THE WIRELESS NETWORK (WN) LAB MISSION IS TO INVESTIGATE, DEVELOP AND PROTOTYPE INNOVATIVE SOLUTIONS IN CYBER-PHYSICAL SYSTEMS FIELD, BUT THE DEVELOPED TECHNOLOGIES HAVE MUCH WIDER APPLICATION OPPORTUNITIES. WN LAB FOCUS HIS WORK IN THE DESIGN OF NEW ALGORITHMS AND TOOLS RANGING FROM THE SMART CITIES TO INTELLIGENT TRANSPORTATION SYSTEMS, AND FROM AMBIENT ASSISTED LIVING SYSTEMS TO MONITORING SYSTEMS.

THE FORMAL METHODS AND TOOLS (FMT) LAB HAS LONGSTANDING EXPERIENCE IN THE DEVELOPMENT AND APPLICATION OF FORMAL NOTATIONS, METHODS, AND SOFTWARE SUPPORT TOOLS FOR THE SPECIFICATION, DESIGN, AND VERIFICATION OF COMPLEX SYSTEMS. THE FMT LAB IS ALSO ACTIVE IN REQUIREMENTS ENGINEERING AND IN VARIOUS BRANCHES OF SOFTWARE ENGINEERING, AND IN THE AREAS OF COMPUTER ETHICS AND INFORMATION TECHNOLOGY AND SOCIETY.

THE SOFTWARE ENGINEERING AND DEPENDABLE COMPUTING (SEDC) LAB DEVELOPS NOVEL INTEGRATED METHODOLOGIES, ROOTED ON THE SYNERGY BETWEEN SOFTWARE ENGINEERING AND DEPENDABLE COMPUTING DISCIPLINES, TO FACE THE CONTINUOUS EVOLUTION AND RISING CRITICALITY OF SOFTWARE-INTENSIVE SYSTEMS.

THE SYSTEM AND SOFTWARE EVALUATION CENTER (SSE) LAB PERFORMS THIRD-PARTY EVALUATION AND CERTIFICATION OF PROCESSES AND PRODUCTS IN THE AREA OF INFORMATION TECHNOLOGY, ACCORDING TO GIVEN REQUIREMENTS AND STANDARDS TO MEET THE NEEDS OF USERS, INDUSTRY AND PUBLIC ADMINISTRATION. INDUSTRIAL CHALLENGES, DEMANDS AND NEEDS IN SYSTEM AND SOFTWARE DEVELOPMENT PROCESSES, ARE THE DRIVERS OF THE SSEC RESEARCH ACTIVITY.

THE ARTIFICIAL INTELLIGENCE FOR MEDIA AND HUMANITIES (AIMH) LAB HAS THE MISSION TO INVESTIGATE AND ADVANCE THE STATE OF THE ART IN THE ARTIFICIAL INTELLIGENCE FIELD, SPECIFICALLY ADDRESSING APPLICATIONS TO DIGITAL MEDIA AND DIGITAL HUMANITIES, AND TAKING ALSO INTO ACCOUNT ISSUES RELATED TO SCALABILITY.

THE HUMAN INTERFACES IN INFORMATION SYSTEMS (HIIS) LAB RESEARCH ACTIVITY AIMS TO ADDRESS FUNDAMENTAL QUESTIONS ON THE INTERACTION BETWEEN PEOPLE AND TECHNOLOGIES AND FOCUSES ON METHODS AND TOOLS TO SUPPORT USER INTERFACE DESIGNERS, SOFTWARE DEVELOPERS, AND END USERS IN OBTAINING SYSTEMS THAT CAN BE ACCESSSED FROM ANY CONTEXT OF USE IN SUCH A WAY TO IMPROVE USABILITY, ACCESSIBILITY, AND USER EXPERIENCE.

THE INFRASTRUCTURES FOR SCIENCE (INFRA SCIENCE) LAB MISSION IS TO CONTRIBUTE TO THE EVOLUTION OF SCIENCE PRACTICES BY INVESTIGATING, EXPERIMENTING, AND CLOSELY CONNECTING RESEARCH AND DEVELOPMENT OF INNOVATIVE DIGITAL INFRASTRUCTURES, INFORMATION SYSTEMS, AND SMART SOLUTIONS FOR FOSTERING AND EMPOWERING DATA-CENTERED RESEARCH.

THE KNOWLEDGE DISCOVERY AND DATA MINING (KDD) LAB IS A JOINT RESEARCH INITIATIVE OF ISTI INSTITUTE OF CNR AND THE DEPARTMENT OF COMPUTER SCIENCE OF THE UNIVERSITY OF PISA. THE OBJECTIVE OF THERE RESEARCH UNIT IS THE DEVELOPMENT OF THEORY, TECHNIQUES AND SYSTEMS FOR EXTRACTING AND DELIVERING USEFUL KNOWLEDGE OUT OF LARGE MASSES OF DATA.

THE HIGH PERFORMANCE COMPUTING (HPC) LAB CARRIES-OUT RESEARCH ON INFORMATION INDEXING AND RETRIEVAL, LARGE-SCALE DISTRIBUTED/CLOUD SYSTEMS, BIG DATA ANALYTICS, MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE, MOBILITY ANALYSIS AND SEMANTIC ENRICHMENT.

THE SIGNALS & IMAGES (SI) LAB GOAL IS TO INCREASE THE KNOWLEDGE IN THE FIELDS OF SIGNAL PROCESSING, IMAGE UNDERSTANDING, AND ARTIFICIAL VISION, IN BOTH THEORETICAL AND APPLICATIVE CONTEXTS. THIS GOAL WILL BE ACHIEVED BY STUDYING AND DEVELOPING MODELS, COMPUTER-BASED METHODS, INTELLIGENT SYSTEMS AND MACHINES FOR THE FORMATION, ELABORATION, ANALYSIS, AND RECOGNITION OF IMAGES AND SIGNALS, AND THEIR APPLICATION IN THE SOCIETY.

THE VISUAL COMPUTING (VC) LAB FOCUS HIS WORK ON THE DESIGN OF NEW COMPUTER GRAPHICS TECHNIQUES, WITH A SPECIAL FOCUS ON GEOMETRY PROCESSING, LEARNING, ACQUISITION, VISUALIZATION, AND PHYSICAL REPRODUCTION. THE VC LAB APPLICATION FIELDS INCLUDE CULTURAL HERITAGE, ENVIRONMENTAL MONITORING, DESIGN, AND ARCHITECTURE.

THE MECHANICS OF MATERIALS AND STRUCTURES (MMS) LAB CARRIES-OUT RESEARCH, SOFTWARE DEVELOPMENT AND CONSULTING IN THE FIELD OF CONTINUUM MECHANICS, WITH PARTICULAR FOCUS ON STRUCTURAL ENGINEERING. MAIN RESEARCH TOPICS ARE: COMPUTATIONAL SOLID MECHANICS, MECHANICS OF MASONRY STRUCTURES, STRUCTURAL HEALTH MONITORING OF HISTORIC BUILDINGS.

THE SPACE FLIGHT DYNAMICS (SFD) LAB IS ACTIVE IN THE FIELDS OF ORBITAL DEBRIS MODELING, MITIGATION AND REMEDIATION, REENTRY PREDICTIONS OF UNCONTROLLED SPACECRAFT AND ROCKET BODIES FOR CIVIL PROTECTION APPLICATIONS, SPACE EXPERIMENTS FOR FUNDAMENTAL PHYSICS, MISSION ANALYSIS, INCLUDING FLIGHT OPERATIONAL SUPPORT, AND ASTRODYNAMICS.

NETWORKING

WIRELESS NETWORKS (WN)

SOFTWARE

FORMAL METHODS AND TOOLS (FMT)

SOFTWARE, ENGINEERING & DEPENDABLE COMPUTING (SEDC)

SYSTEM AND SOFTWARE EVALUATION (SSE)

KNOWLEDGE

ARTIFICIAL INTELLIGENCE FOR MEDIA AND HUMANITIES (AIMH)

HUMAN INTERFACES IN INFORMATION SYSTEMS (HIIS)

INFRASTRUCTURES FOR SCIENCE (INFRA SCIENCE)

KNOWLEDGE DISCOVERY AND DATA MINING (KDD)

HIGH PERFORMANCE COMPUTING

HIGH PERFORMANCE COMPUTING (HPC)

VISUAL

SIGNALS AND IMAGES (SI)

VISUAL COMPUTING (VC)

FLIGHT AND STRUCTURAL MECHANICS

MECHANICS OF MATERIALS AND STRUCTURES (MMS)

SPACE FLIGHT DYNAMICS (SFD)



Istituto di Scienza e Tecnologie
dell'Informazione "A. Faedo"
Consiglio Nazionale delle Ricerche



Deepfakes: A Digital Transformation Leads to Misinformation

Nika Nour and Julia Gelfand

University of California, Irvine, USA

Abstract

Deepfakes are a product of artificial intelligence (AI) and software applications used to create convincing falsified audiovisual content. Linguistically, a portmanteau combines deep learning aspects of AI with the doctored or falsified enhancements that deem the content fake and now deepfake or misinformation results. A variety of sophisticated software programs' exacting algorithms create high-quality videos and manipulated audio of people who may not exist, twisting others who do exist, creating the potential for leading to the spread of serious misinformation often with serious consequences. The rate of detection of this digital emergence is proliferating exponentially and the sourcing is challenging to verify, causing alarms. Examples of this pervasive information warfare are is associated with deepfakes that range from identity theft, discrediting public figures and celebrities, cyberbullying, blackmail, threats to national security, personal privacy, intensifying pornography and sexual exploitation, cybersecurity, baiting hate crimes, abusing social media platforms and manipulating metadata.

Deepfakes that are difficult to cite, acquire, or track have some parallel attributes to grey literature by that definition. Often detectable, yet problematic, activities such as phishing and robocalling may be common attempts of deepfake activities that threaten and interrupt rhythms of daily life. The increasing online personas that many people create or assume contribute to this fake content and potential for escalated exploitation due to technical abilities to copy and reimagine details that are not true. AI image generators create completely false images of people that simply don't exist within seconds and are nearly impossible to track. While AI is perceived as a positive benefit for science and policy, it can have negative roles in this new AI threatened environment. Deepfakes have cross-over targets in common business applications and society at large. Examples of this blur are targeted advertising, undetected security cameras in public spaces, blockchain, tabloid press/paparazzi, entertainment, computer games, online publishing, data and privacy, courtroom testimony, public opinion, scientific evidence, political campaigns, and rhetoric.

This paper explores the impact and intersections of these behaviors and activities, products of AI, and emerging technologies with how digital grey and the optics of grey expose the dangers of deepfakes on everyday life. Applying a security and privacy lens, we offer insights of extending libel and slander into more serious criminal behavior as deepfakes become more pervasive, construing reality, endangering personal, social, and global safety nets adding to the new normal we assume today. How we became more sensitized to misinformation and fake news tells the story about deepfakes.

Introduction

Grey Literature (GL) has a history and nomenclature that includes a definition that suggests it is a "relatively recent collective noun" for "information produced on all levels of government, academia, business and industry in electronic and print formats not controlled by commercial publishing where publishing is not the primary activity of the producing body" (Farace and Frantzen 2005). Traditional scholarly literature is defined by having its submissions go through peer review. Often, grey literature is not peer-reviewed as it is not usually published in journal literature or monographs released by commercial or scholarly imprints. This suggests the parallels we see with some deepfake strategies

and products since they too may not be the primary output of their sources and certainly are not peer-reviewed.

Other critical intersections include how findings from the 'Project Overview of the Grey Literature Strategies' concluded best practice guidelines for producing and managing grey literature in Australia to transform access to public interest research for various communities (Aloia and Naughton, 2017). Considering the news trajectories of the past few years and the political and social climate we are currently experiencing, this is a very timely assessment. Part of the Australian research and its outcomes demonstrated that "finding better ways to access, control, evaluate, collect and preserve grey literature is an important national and international issue" (Houghton, 2011) in our current environment, much as it was leading up to today. Publishing strategies have changed as electronic publishing and preservation options have widely expanded the creation and dissemination of information linking different formats such as text, video, data, imagery, and other content into single outputs. Five years ago, it was reported that "Australia produces \$30 billion worth of 'grey literature' that we can't read" (McCallum, 2016), suggesting how government, academe, corporate entities, and other information providers issue content that is not well cited nor curated. The value of these resources is understated due to poor discovery and application. Some examples of new databases containing content from many different repositories and sources are now practicing commendable stewardship, safeguarding, indexing, and curation to protect "endangered" or lost content that is getting recognition and inclusion in contemporary research. One such example is Policy Commons, launched in 2020 by Coherent Digital to be recognized for its aggregation of premium full-text content, subject neutral but research-intensive and described by co-founder Toby Green as "Essential research done by IGOs, NGOs, think tanks and research centers simply doesn't get the attention it deserves. It's hard to find, it disappears when funding is lost, it lacks persistent identifiers, and Policy Commons fixes these issues" (Green, 2020). That is the difference in what defines something as grey.

The study of grey literature has matured over the last couple of decades as its importance has been applied to new fields, disciplines, technologies, products, and outcomes. With that has come the refinements of how specialized content that forms the nexus of grey literature has shown the importance of its applications in many fields such as sciences, technology, health, government operations, public policy, and even arts and visual studies (Pappas and Williams, 2011). The linking factor of communications, and perhaps the intersection of information, communication, and technology best described by ICTs has shown how specialized content of working papers, technical reports, illustrations, images, spatial and quantitative data, legislation, thesis/dissertations, lyrics, product descriptions, plans, blogs, tweets, and hosts of other information outputs has had a dizzying effect on the publishing landscape as each has become increasingly valuable to communicate with readers and informants about certain observations, opinions, conclusions, and other outputs challenging to find, source, cite and share in the everyday lexicon of use. This challenged how it was collected, described, transmitted, and used.

During the Coronavirus/COVID 19 pandemic, the health sciences, and government agencies, in particular, were faced with trying to convey the medical consequences of the pandemic and direct citizens to act responsibly and rationally in responding to this worldwide crisis. However, efforts at communicating how best to practice recommendations and guidelines were met with resistance. For instance, the public questioned mandates and practices to honor social distancing, personal hygiene practices, engage in tracing and testing, and be vaccinated once shots became available (Brennan, 2020). The effort to focus on scientific evidence rather than conspiracy theories

and opinions that lacked medical data has been challenging, framing one of the best examples of how misinformation has flourished by different media establishments and outlets over the past two years testing credibility and legitimacy at every turn. One of the most prominent examples of COVID-19 misinformation is the long list, known as a factsheet compiled by Brennen et al. where this team analyzed a sample of 225 pieces of misinformation rated false or misleading by fact-checkers and published in English between January through March 2020, drawn from a collection of fact checks maintained by the British First Draft News (Brennen et al., 2020). Social media has contributed to both the good and negative information sharing that defines our knowledge base today.

Understanding Misinformation

We have adopted the following definitions by Wardle and Derakhshan and share their sentiments that “it’s important to distinguish true and false messages, as well as messages that are and are not created, produced or distributed with intent” (Wardle and Derakhshan, 2019).

- Misinformation - false information shared by someone who believes to be true
- Disinformation - by contrast, is false information shared with knowledge of its falsity and thus intention to deceive or otherwise do harm. It is a deliberate, intentional lie.
- Malinformation - information based on a reality that is shared to harm a person, organization, or country. This term can refer to instances where private information is made public or genuine imagery is reshared in the wrong context.
- Information disorder - an umbrella term encompassing all forms of the above

Grey Literature has not been accused of serving up misinformation or fake news; however, it has been aligned with questions about why this sourcing is so fragile and why such valuable content has fallen through the cracks in what is considered a sophisticated publishing network. Several reasons begin to tell this tale. Local, federal, and international government agencies and nonprofit organizations have been forced to make critical decisions about what to archive and how to safeguard and distribute it (Langa, 2021). The financial side of the publishing enterprise is increasingly volatile as the call for open access has been made loud and clear in academic and broader circles (Fraga-Lamas and Fernandez-Carames, 2020). As “born digital” became more common, the release and citation of this material became less structured, and copyright protections were reduced. According to Suzanne Smalley, Roger Schonfeld speculates “how misinformation, politicization and other problems embedded in the open-access movement stem from a “mismatch” between the incentives in science and how ‘openness and politicization are bringing science into the public discourse” (Schonfeld, 2021).

The Good, The Bad and the Future of both Grey Literature and Deepfakes

As Richard Van Hooijdonk states in his blog, “The competition between the creation and elimination of deepfakes will become increasingly fierce in the future, with deepfake technology not only becoming easier to access but its content easier to create and harder to distinguish from real” (Van Hooijdonk, R, 2021). This is not the case with grey literature as it more uniformly has reduced the gap between traditional commercial publishing and other new forms of expression and types of content, extending the open access movement. It can be perceived as Van Hooijdonk suggests that “the ability to use artificial intelligence to create realistic simulations might even be a positive thing for humanity” (2021).

Grey literature has been deeply seeded in medical and scientific areas from its inception. This transfer can also be seen in the correlated history of the open access movement when the demand for access to medical information was made by patients

who had no systematic access to information about medical conditions, diagnoses, and treatments their physicians and healthcare providers were recommending as most of the information was published in subscription-based journals held by special libraries to which the unaffiliated public was not allowed.

Wardle and Derakhshan's "Misinformation and Disinformation Framework" when applied to documentary films and media and what challenges they pose to audiences and perhaps readers is yet another example of what they offer as a new lexicon to help to distinguish "between media commentary, misinterpreted material, playful content and media created to deliberately mislead" (Hight, 2021). Summarized below are several examples of different forms captured by Hight from Wardle's framework that we find insightful when describing how misinformation is created:

- Fabricated content – deliberately designed to deceive, completely false
- Manipulated content – where there has been a manipulation of genuine material;
- Imposter content – when genuine sources are impersonated;
- False content – genuine content mixed with false contextual information;
- Misleading content – misleading use of information to frame an issue or individual;
- False connection – when headlines, visuals, or captions don't support the content;
- Satire of parody – fake content intended for social commentary

In addition, Claire Wardle has extended content in her created definitional toolbox, offering a glossary. She maps the landscape and describes graphically more about Information Disorders indicating how these tools will help contextualize and categorize how we can more appropriately address issues of trust and truth in the digital age (Wardle, 2018). She contributes a bright assessment about how these categories should be used in parallel and should be commended for the work that she has contributed to *First Draft*, where she has created a collaboration to "stand up for truth in a polarized world" (Wardle, 2018).

Another lesson about misinformation relevant to grey literature is how Nossel addresses disinformation containment strategies. She says "the originators of disinformation – not just foreign governments but conspiracists, provocateurs, and paid propagandists – are too diffuse to be shut down; even trying to shut them down would unavoidably impinge on expressive rights." She continues her drift with "Stopping disinformation will also require a more refined understanding of who consumes it and why" and concludes that information consumers can be divided into the anchored, the adrift and the marooned." (Nossel, 2021). This suggests our agreement that it is the middle group, the "informationally adrift," as Nossel refers to them, who are most challenging to educate or inform because of their over exposure to content, they are "prone to perpetual doubt" and are part of growing communities of disbelief and denial. This means that they can't distinguish sourcing and determine what is rightful or credible. The challenge is to urge this diverse body to do their research in trusted sources building media literacy and grey literature when made available, archived and heavily cited contributes to that slow fix while parsing misinformation and disinformation.

With the quickly changing media landscape and generations of users relying on multiple social media platforms, findings from Aswani and colleagues are equally alarming and support that misinformation is created by misinformation propagators, and how it is being shared. Their Twitter analysis found that "misinformation was 43% of tweets were works of fiction, 27% were rumors and 22% was from vested interests of organizations and individuals for the purpose of content promotion and advertising with only 8% from government, politicians and media sources combined." (Aswani, et al, 2019).

Misinformation Leading to Deepfakes

Even though artificial intelligence and online misinformation consistently make headlines, minimal research exists on digitally-altered visual content such as AI-generated deepfake videos. This paper focuses on video experience with deepfakes rather than covering the entire spectrum of deepfake creation. Using deep learning algorithms, “deepfake” videos (or colloquially, just “deepfakes”) typically substitute one person’s visual and acoustic likeness for another, presenting viewers with compelling videos of individuals doing and saying things they never did or said (Vaccari et al., 2020). In addition, many politicians and media pundits believe that deepfake videos can influence elections, instigate violence, and destroy claims to the truth (Paris et al., 2010). Although verifying online content and imagery is not a new phenomenon—in fact, the entire field of image forensics exists for this purpose—academic research exploring this topic is only recently emerging. However, due to the rising prevalence of deepfake videos and their ability to promulgate violence and mistrust, it is increasingly vital for academics to deploy their technical and theoretical expertise in combating misinformation online.

What makes deepfake videos especially worrisome is the relative ease and accessibility with which adverse actors can manipulate moving images. With virtually no technical expertise, individuals can produce untraceable, deceptive videos and distribute them online from almost anywhere in the world. Altered videos have the power to propel a terrorist group’s agenda or reword a politician’s speech. The critical factors for assessing non-textual media distribution are the rapid pace of technology, the recent growth in digitally altered content, and users’ varying levels of trust in visual imagery. While algorithmic detection for deepfake videos exists, these technologies’ access and deployment remain uneven, and artificial intelligence blockers have struggled to eliminate harmful online posts and publications (Delort and Paris, 2011). Historically, this type of conduct has led to name-calling tactics, releasing personally identifiable information (doxing), and similar intentions to create reputational harm. Deepfake videos, armed with the rhetorical persuasiveness of the moving image, represent a significant jump in the ability to inflict social damage.

Acknowledging that the algorithmic detection of deepfake videos is mired in its troubles of reliability, accessibility, and credulity, there is a robust, practical impetus to evaluate—and, where necessary—cultivate the human capacity for discerning between legitimate and deepfake manipulated videos (Biometric, 2019). In this paper, we document the process of creating deepfake videos as research instruments to test the ability of individuals to detect deepfakes in a controlled setting, outlining both the steps taken and the methodological motivations for them. We also draw comparisons to other information outputs. Finally, our documentation suggests that multiple considerations must be taken into account when creating altered content for research purposes, particularly for the sake of generalizability, concerning how individuals encounter altered content online.

The Rise and Ethics of Deepfakes

Manipulating visual media is becoming widely accessible, inexpensive, and easier to accomplish (Wade, et al, 2002). With software becoming more user-friendly, cell phone cameras increasing in quality, and the rise of technology literacy, most people have the opportunity to doctor media, an option once reserved for video editing professionals and filmmakers. Fake news and artificial intelligence consistently make headlines in today's media market; however, there is minimal research on digitally altered, visual content, specifically deepfake videos. As deepfakes rise in prevalence, there are concerns about the medium's ability to influence critical societal functions and political outcomes. Even memory researchers are studying how falsified videos can plant memories through familiarity, imagery, and credibility in how these manipulated content pieces are presented (Nash, Wade, and Brewmann, 2009). Without developing detection strategies, deepfakes can invoke personal and societal harm, threatening the foundations of trust and society. As the internet user's confidence in content fades and false information is presented as accurate, society becomes more susceptible to information warfare.

Algorithmic Detection

While some scholarly reviews of deepfake algorithms like FakeApp, Adobe VoCo, Lyrebird, and Face2Face have tended towards the ethically agnostic (Gardiner 2019), other recent research has shifted towards the creation of methods to detect deepfake manipulations, with both the explicit and implicit standpoints that deepfake videos pose social and moral dangers to the general public. Deep learning and machine learning feature prominently in the production of algorithmic attempts to detect deepfakes and—by extension of their detection—to impune on their persuasiveness of deepfake videos distributed online.

Image forensics research in the area of deepfake videos has focused on the use of neural networks (Güera and Delp, 2018; Amerini et al., 2019; Guarnera et al., 2020), classical frequency domain analysis (Durall et al., 2019), facial recognition (Korshunov and Marcel, 2019), and the algorithmic detection of visual artifacts that emerge during deepfake manipulation (Marra et al., 2018; Li et al., 2019). Problematically, as observed by Nguyen et al. (2019), the very technologies used to detect deepfake videos are the same as those used to create deepfake videos, thus perpetuating an arms race hinged on the co-advancement of shared algorithms.

Human Detection

Research on human (as opposed to machine) deepfake detection has more strongly figured around issues of literacy (Nightingale et al., 2017; Schetinger et al., 2017) and the social outcomes of deepfake circulation across social media platforms (Vaccari et al., 2020). Given that the creation of deepfake forensic algorithms consequently advances the ability of deepfake video creators to evade algorithmic detection further, there is an urgent need to cultivate and advance the human detection of deepfake manipulations. Moreover, the circumstances in which individuals encounter deepfake videos unwittingly differ dramatically from the large, high-resolution datasets that algorithmic detection models learn from. For instance, in Durall et al.'s work, as image resolution for deepfake videos fell, so too did their algorithm's predictive accuracy, suggesting that we cannot entirely rely on machines to detect manipulations in low-resolution videos correctly (2019). When coupled with Marra et al.'s observations on how the circulation of deepfake videos occurs primarily via social media channels that compress images and videos—and in doing so, obfuscate algorithmic detection "signals" with low-resolution, artificial "noise"—it becomes increasingly evident that the need for human detection further

increases (particularly when access to algorithmic detection models is uneven, impractical, and inequitable (2018).

Additional Applications

Other related technologies like drones that follow and track individuals give the public some of the same insecurity since they have not authorized or given permission to be followed or tracked. However, we can't forget how through the 2020 U.S. Presidential Election news outlets provided fact-checking after nearly every presidential debate and large campaign rally trying to force candidates to substantiate their claims with accurate history, chronology, and sourcing, rather than high-risk cases of any of the above forms of misinformation.

Data as an example of grey literature is reflected by Polonetsky, Tene, and Finch in their article, "Shades of Gray: Seeing the Full Spectrum of Practical Data De-Identification" (2016). They propose "parameters for calibrating legal rules to data depending on multiple graduations of identifiability, while also assessing other factors such as an organization's safeguards and controls, as well as the data's sensitivity, accessibility, and permanence" (Polonetsky et al., 2016). "It builds on emerging scholarship that suggests that rather than treat data as a black or white dichotomy, policymakers should view data in various shades of gray; and provides guidance on where to place important legal and technical boundaries between categories of identifiability." (Polonetsky et al., 2016, 595). They go on to create a data spectrum of key inflection points that include:

- Explicitly Personal Data
- Potentially Identifiable & Not Readily Identifiable Data
- Key-coded Data
- Pseudonymous and Protected Pseudonymous Data
- De-Identified and protected De-Identified Data
- Anonymous and Aggregated Anonymous Data

Their work continues to justify how important it is to separate between the sensitivity of a data item and its degree of identifiability and concludes with "New uses of data and technology have the potential to bring humanity a wide range of benefits, but at the same time to generate new and serious harms" while advancing "an approach that supports benefit and deters risk by providing a practical framework for policymakers to analyze various data sets based on their degree of identifiability" (Polonetsky et al., 2016, 623).

Reflections of Technology

The Evolution of Media Manipulation

Manipulating photos and videos isn't a new concept. Instead, an early example of media manipulation includes a portrait of Abraham Lincoln in 1860. Though the image seemed real, it was actually multiple photographs of Lincoln's head and John Calhoun's body stitched together into one portrait (Caldera, 2020). As image and photo manipulation became more common, video editing and manipulated also took hold as early as the 1970s where computer animation and visual effects became a widely adopted technique in the entertainment industry (Caldera, 2020). However, the rise of hyper-realistic simulations and media have far surpassed the simple edits made to videos and photos and taken a much more nefarious turn in recent years. As media technologies advance and develop synthetic results, falsified audiovisual content is appearing and sounding more realistic (Johnson and Diakopoulos, 2021). Examples of the deepfake

phenomenon include videos celebrities doctored into movie clips they never performed to Mark Zuckerberg stating that he was going to delete Facebook (Kietzmann et al., 2019). Though falsifying images, audio, and video content isn't a new concept, the phenomenon of deepfakes led to a boom of this new type of false media when an anonymous Reddit user shared the computer code to place famous personalities into pornographic clips (Kietzmann et al., 2019). The accessibility and feasibility of deepfakes have also rapidly evolved. Today's deepfakes are made with minimal coding knowledge and require no elaborate hardware. Advanced computer algorithms and apps can quickly generate, produce, and edit videos that are difficult to tell apart from the original content (Skibba, 2020). As deepfakes become more prevalent on social media platforms and mobile networks, generated adversarial networks (GANs) allow for faster dissemination and create vulnerabilities for face recognition software (Korshunov and Marcel, 2018). With facial recognition becoming less reliable in these circumstances, this technology poses a new threat to the foundation of trust online as people may be unable to detect or determine whether or not these videos are real. While some of the existing deepfake videos seem harmless, the potential consequences of this technology can impact citizenry, social welfare, business, and life in general.

The Consequences of Deepfakes in Everyday Life

As fake content becomes more prolific in the digital age, we must understand the current and potential societal impact of deepfakes on the citizenry, education, social welfare, politics, media, business, and family life. The creation of deepfake videos is becoming easier and more accessible for unskilled end-users through programs such as FakeApp, a face-swapping program, Zao, a Chinese mobile app using movie clips, and Apple's text-to-speech (TTS) editing system (Kietzmann, 2019).

Accessibility and new technologies suggest how much deepfakes are increasingly grey. Superimposing and swapping faces of individuals into movies and television clips can be an entertaining example of deepfake videos. For instance, in 2019, a clickbait video entitled "Keanu Reeves Stop A ROBBERY " included a stuntman and voice actor with his face replaced with celebrity actor Keanu Reeves (Bode, 2021). The video was shared across multiple social media platforms, but the content was falsified, and Keanu Reeves was never involved. Though this type of content can be amusing, the ease of deepfake content creation means accountability is lacking, and entertainment becomes nefarious information dissemination. The fact that such changes are not chronicled provides no record of artifacts or tracing available. Anyone, anywhere, at any time, can create these doctored videos to convince the general public that they are real. However, the ability to track, hold accountable, or impose regulations on these perpetrators are practically nonexistent or require legal and financial resources often dismissed. For instance, with limited oversight, deepfakes can undermine trust in the news and elections, resulting in ethical implications (Diakopoulos and Johnson, 2019). In 2017, researchers successfully created deepfake stills of former President Barack Obama saying things that he never actually said but were voiced by actor, Jordan Peele (Citron and Chesney, 2019). The same technology can create false videos inciting plans to carry out political assassinations, seemingly private conversations featuring elected leaders (Citron and Chesney, 2019), or Speaker Nancy Pelosi slurring her speech (Denham, 2020). With deepfakes being the next frontier for fraud, companies like Recorded Future found multiple examples on the dark web of criminal activity using deepfakes to blackmail, create pornographic videos, and execute identity theft (Security Firm, 2021). Society and government entities need appropriate countermeasures when such actions impact

personal liabilities such as non-consenting individuals, government officials, and organizations (Kietzmann et al., 2019).

Conclusions

Our exploration of this topic moves us to share the conclusion stated by Vizoso, et al, that deep fakes are highly understudied and how they will continue contributing to more misinformation will only proliferate. We speculate that there are many new research paths that can emerge as we observe how journalists, media creators, information analysts and others respond to the damage that deep fakes can cause to society and how technology and media literacy can potentially change that course. (Vizoso et al, 2021).

Grey literature may be reduced as more outputs become discoverable. As search, meta-data, and analytics become more embedded in digital materials, deepfakes are newer, multi-formatted, and potentially more prolific. With terrorism, privacy, distribution channels, relationships to blockchain, and cryptocurrency, information warfare is regularly redefined and more concerning. This leads to an unstable digital marketplace where even the definition of grey is evolving. These global concerns come with implications as stakeholders and authorities enter the wild west of policy and legal oversight. These unknown territories will only expand and cross digital borders as deepfake creators become more aggressive and confident in their abilities to create harm. The scale at which these moving targets are changing will dictate how regulatory bodies respond and develop legal frameworks in this grey area of policing the digital arena. Digital transformation is an ongoing process, especially in the space where deepfakes are proprietary but not identifiable, definable, or attributable. As society learns to govern and respond to the new grey, individuals must take it upon themselves to build media literacy and resiliency when faced with all forms of information disorder.

References

- Aloia, D. and Naughton, R. (2017). The Greylit Report: Understanding the challenges of finding grey literature. *Grey Journal*. 13. 75-80. <https://doi.org/10.17026/dans-2z8-x27y>
- Amerini, I., Galteri, L., Caldelli, R. and Del Bimbo, A. "Deepfake Video Detection through Optical Flow based CNN," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.
- Aswani, R., Kar, A.K., and Ilvarasan, P.V. (2019). Experience: Managing Misinformation in Social Media – Insights for Policymakers from Twitter Analytics. *ACM Journal of Data and Information Quality* 12 (1), Article 6. <https://doi.org/10.1145/3341107>
- Bode, L. (2021). Deepfaking Keanu: YouTube deepfakes, platform visual effects, and the complexity of reception. *Convergence: The International Journal of Research into New Media Technologies* 27 (4) 135485652110304-934. <https://doi.org/10.1177/13548565211030454>
- Brennen, J.S., Simon, F.M., Howard, P.N., and Nielsen, R.K. (2020). Types, Sources and Claims of COVID-19 Misinformation.
- Caldera, Elizabeth (2020) " 'Reject the Evidence of Your Eyes and Ears': Deepfakes and the Law of Virtual Replicants," *Seton Hall Law Review*: Vol. 50 : Iss. 1 , Article 5. Available at: <https://scholarship.shu.edu/shlr/vol50/iss1/5>
- Chesney, Robert and Danielle Citron, (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy and National Security. *California Law Review* 107.
- Citron, D.K. and Chesney, R. (2019). Deepfakes and the New Disinformation War. *Foreign Affairs*. Available at: https://scholarship.law.bu.edu/shorter_works/76
- Coherent Digital (2020, November 2). Press Release: *World's largest resource for public policy launches today*. Retrieved November 5, 2021 from <https://coherentdigital.net/policycommonslaunch>
- Deepfake videos easily fool face systems, researchers warn. (2019). *Biometric Technology Today*. (10), 3–3. [https://doi.org/10.1016/s0969-4765\(19\)30137-7](https://doi.org/10.1016/s0969-4765(19)30137-7)

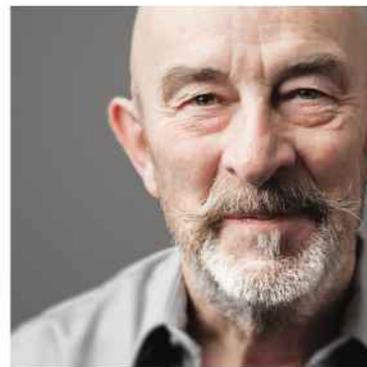
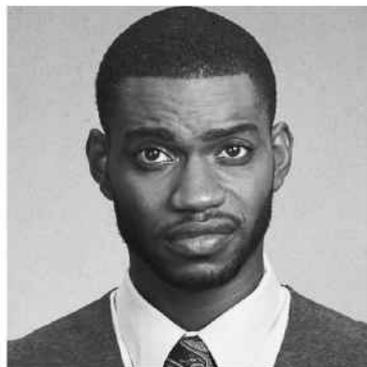
- Delort, J., Arunasalam, B., and Paris, C. (2011). Automatic Moderation of Online Discussion Sites. *International Journal of Electronic Commerce*, 15 (3), 9-30. doi:10.2753/jec1086-4415150302
- Denham, H. (2020, August 3). *Another fake video of Pelosi goes viral on Facebook*. The Washington Post. Retrieved December 2, 2021 from <https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fake-video-facebook/>.
- Diakopoulos, N. and Johnson, D. (2019). Anticipating and Addressing the Ethical Implications of Deepfakes in the Context of Elections. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3474183>
- Durall, R., Keuper, M., Pfreundt, F., and Keuper, J. (2019). Unmasking DeepFakes with simple Features. <https://arxiv.org/abs/1911.00686>
- Farace, D., Frantzen, J., Schöpfel, J., Stock, C. and Boekhorst, A. (2006). Access to Grey Content: An Analysis of Grey Literature based on Citation and Survey Data: A Follow-up Study. In: Seventh International Conference on Grey Literature: Open Access to Grey Resources, Nancy, France, December 5-6, 2005. - *GL7 Conference Proceedings*, 194-203.
- Fraga-Lamas, P. and Fernández-Caramés, T. (2020). Fake News, Disinformation, and Deepfakes: Leveraging Distributed Ledger Technologies and Blockchain to Combat Digital Deception and Counterfeit Reality. *IT Professional*. 22. 53-59. <https://doi.org/10.1109/MITP.2020.2977589>
- Gardiner, N.J. (2019). *Facial re-enactment, speech synthesis and the rise of the Deepfake*. https://ro.ecu.edu.au/theses_hons/1530
- Gelfand, J. and Lin, A. (2017). Social Media Matters: Showing Up Online as Well as Ontime. Paper presented at the 19th International Conference on Grey Literature. Rome, Italy. October 23, 2017. <https://av.tib.eu/media/37267>
https://greyguide.isti.cnr.it/attachments/category/34/Gelfand_and_lin.pdf
- Gelfand, J. and Tsang, D. (2015). Data: Is it Grey, Maligned or Malignant? Paper presented at the 16th International Conference on Grey Literature, Washington, DC, December 9, 2014. *The Grey Journal*. 11(1), 2015: 30-40.
- Green, T. (2020). World's Largest Resource for Public Policy Launches Today. Press Release by Coherent Digital, November 2. <https://coherentdigital.net/policycommonslaunch>
- Guarnera, L., Giudice, O., & Battiato, S. (2020). DeepFake Detection by Analyzing Convolutional Traces. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2841-2850.
- Guera, David & Delp, Edward. (2018). Deepfake Video Detection Using Recurrent Neural Networks. 1-6. 10.1109/AVSS.2018.8639163.
- Hight, C. (2021). Deepfakes and Documentary Practice in an Age of Misinformation. Continuum, DOI: [10.1080/10304312.2021.2003756](https://doi.org/10.1080/10304312.2021.2003756)
- Johnson, D. G. and Diakopoulos, N. (2021). What to do about deepfakes. *Communications of the ACM*. 64. 33-35. <https://doi.org/10.1145/3447255>
- Kietzmann, J., Lee, L., McCarthy, I., and Kietzmann, T. (2019). Deepfakes: Trick or treat? *Business Horizons*. 63. <https://doi.org/10.1016/j.bushor.2019.11.006>
- Korshunov, P., & Marcel, S. (2018). DeepFakes: A New Threat to Face Recognition? Assessment and Detection. *ArXiv, abs/1812.08685*.
- Langa, J. (2021). Deepfakes, Real Consequences: Crafting Legislation to Combat Threats Posed by Deepfakes. *Boston University Law Review* 101(2), 761-801.
- Lawrence, A., Houghton, J., Thomas, J. and Weldon, P. (2014). Where Is the Evidence? Realising the Value of Grey Literature for Public Policy & Practice, A Discussion Paper. <https://apo.org.au/node/42299>
- Marra, F., Gagnaniello, D., Cozzolino, D., and Verdoliva, L. Detection of GAN-Generated Fake Images over Social Networks. (2018). 384-389. <https://doi.org/10.1109/MIPR.2018.00084>
- McCallum, T. (2016). "How Australia Produces \$30 Billion Worth of 'Grey Literature' That We Can't Read." *The Conversation*, April 27, 2016. Retrieved November 11, 2021 from <https://theconversation.com/how-australia-produces-30-billion-worth-of-grey-literature-that-we-cant-read-56584>
- McGann, James. (2020). *2020 Global Go to Think Tank Index Report*. https://repository.upenn.edu/think_tanks/18/
- Mercer, K., Weaver, K.D., and Waked, K. (2021). Navigating Complex Authorities: Intellectual Freedom, Information Literacy and Truth in Pandemic Stem Information. *IFLA Journal*, <https://doi.org/10.1177/03400352211048915>
- Mirsky, Y. and Lee, W. (2020). The Creation and Detection of Deepfakes: A Survey. *ACM Computing Surveys*, 54 (1), Article 7. <https://doi.org/10.1145/3425780>
- Nash, R. A., Wade, K. A., & Brewer, R. J. (2009). Why Do Doctored Images Distort Memory? *Consciousness and Cognition*, 18(3), 773-780. <https://doi.org/10.1016/j.concog.2009.04.011>
- Nightingale, S. J., Wade, K. A., & Watson, D. G. (2017). Can People Identify Original and Manipulated Photos of Real-World Scenes? *Cognitive Research: Principles and Implications*, 2(1), 30. <https://doi.org/10.1186/s41235-017-0067-2>

- Nguyen, T.T., Nguyen, C.M., Nguyen, D., Nguyen, D.T., and Nahavandi, S. (2019). Deep Learning for Deepfakes Creation and Detection. <https://arxiv.org/abs/1909.11573>
- Nossel, S. (2021). How to Save People from Drowning in a Sea of Misinformation. *Slate*, December 15. https://slate.com/technology/2021/12/information-consumers-misinformation-adrift-media-literacy.html?fbclid=IwAR3REIsINufZrqM29U_FEDIOorTdYXH-T8x6BkHLzIzqM4HoraB52v6G-x8
- Pappas, C. and Williams, I. (2011). Grey Literature: Its Emerging Importance, *Journal of Hospital Librarianship*, 11(3), 228-234, <https://doi.org/10.1080/15323269.2011.587100>
- Paris, Britt, and Donovan, Joan. (2019). Deepfakes and Cheap Fakes. *Data & Society*. <https://datasociety.net/library/deepfakes-and-cheap-fakes/>
- Polensky, J., Tenem, O., and Finch, K. (2016). Shades of Gray: Seeing the Full Spectrum of Practical Data De-identification." *Santa Clara Law Review*, 56 (3), 593-628.
- Policy Commons*. (n.d.). Retrieved December 2, 2021 from <https://policycommons.net/>
- Saunders, L. and Budd, J. (2020). Examining, Authority and Reclaiming Expertise. *Journal of Academic Librarianship* 46, 102077. <https://doi.org/10.1016/j.acalib.2019.102077>
- Schetinger, Victor & Oliveira, Manuel & Silva, Roberto & Carvalho, Tiago. (2015). Humans Are Easily Fooled by Digital Images. *Computers & Graphics*, 68. <https://doi.org/10.1016/j.cag.2017.08.010>
- Schonfeld, R. (2021). Is Scientific Communication Fit for Purpose? *Scholarly Kitchen*, November 1. Retrieved from <https://scholarlykitchen.sspnet.org/2021/11/01/is-scientific-communication-fit-for-purpose/>
- Security firm: Deepfakes are 'fraud's next frontier.' (2021). *Biometric Technology Today*, (6), 2-3. [https://doi.org/10.1016/s0969-4765\(21\)00064-3](https://doi.org/10.1016/s0969-4765(21)00064-3)
- Skibba, Ramin. (2020). Accuracy Eludes Competitors in Facebook Deepfake Detection Challenge. *Engineering*, 6. 1339-1340. <https://doi.org/10.1016/j.eng.2020.10.008>
- Smalley, S. (2021, April 13). As Misinformation Grows, Scholars Debate How to Improve Open Access. Inside Higher Ed. Retrieved from <https://www.insidehighered.com/news/2021/11/08/open-access-science-misinformation-era>
- Vaccari, C. and Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*. 6. 205630512090340. <https://doi.org/10.1177/2056305120903408>
- Van Hooijdonk, R. (2021, April 23). The Good, the Bad, and the Future of Deepfakes [web log]. Retrieved November 11, 2021 from <https://richardvanhooijdonk.com/>
- Vizoso, A. Vaz-Alvarez, M. and Lopez-Garcia, X. (2021). Fighting Deepfakes: Media and Internet Giants' Converging and Diverging Strategies Against Hi-Tech Misinformation. *Media and Communication*, 9 (1): 291-300. <https://doi.org/10.17645/mac.v9i1.3494>
- Wade, K., Garry, M., Read, J., and Lindsay, D. (2002). A Picture is Worth a Thousand Lies: Using False Photographs to Create False Childhood Memories. *Psychonomic Bulletin & Review*. 9. 597-603. <https://doi.org/10.3758/BF03196318>
- Wardle, C. (2018, July 9). *Information disorder: The definitional toolbox*. First Draft. Retrieved November 2, 2021 from <https://firstdraftnews.org/articles/infodisorder-definitional-toolbox/>
- Wardle, C. (2019). First Draft's Essential Guide to Understanding Information Disorder." *First Draft News*. Retrieved November 11, 2021 from https://firstdraftnews.org/wpcontent/uploads/2019/10/Information_Disorder
- Zimdars, M. and K. McLeod, eds. (2020). *Fake News: Understanding Media and Misinformation in the Digital Age*. Cambridge, MA: MIT Press, 2020.



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

Grey Literature citation and inclusion rates in gambling review articles: Opportunities for improvement

David Baxter

Department of Political Science, University of Alberta

Abstract

Commercial gambling is expanding in many countries worldwide, and researchers and governments are increasingly approaching gambling as a public health issue. Grey literature is a popular avenue for disseminating findings from gambling studies, accounting for over 20% of gambling research publications. As evidence-based policy decisions increasingly rely on systematic and scoping reviews, it is important that grey literature evidence is included in these reviews. To date, two umbrella review has assessed systematic reviews for gambling-related interventions, and using the AMSTAR 2 critical appraisal tool, both found the overall quality of the reviews to be low. One of the AMSTAR 2 criteria is “a comprehensive literature search strategy”, which includes grey literature where relevant. The goal of this study is to assess the extent to which grey literature is cited in gambling related review articles, how often grey literature searches are included in reviews with systematized search strategies, and how grey literature is discussed (or not).

Using the multidisciplinary scholarly database Web of Science, a broad keyword search for “gambl” was performed, limited to documents of type “Review” published from 2016 to 2020. After screening for articles unrelated to gambling, 174 reviews were included. The references section of each article was reviewed and all grey references were catalogued. For articles with a formal search strategy, the proportion of grey items included in the systematic search results was also recorded. To determine if and how grey literature is discussed, the full text was searched for grey literature related terms.*

Of the 174 included studies, 100 had systematized search strategies and 74 did not. In systematized reviews, grey literature sources were included in just over half of reviews (n=54). Of the 46 systematized reviews that excluded grey literature, only one provided a methodologically sound reason for doing so and only eight acknowledged grey literature exclusion as a limitation of the review, while many did not mention grey literature at all. Across all review types, grey literature represented an average of 9.2% of works cited. Compared to similar reviews in other domains, it appears that grey literature is underutilized in gambling reviews. Efforts to improve grey literature uptake in gambling reviews may be most effective if focused on increasing awareness of grey literature and providing information or training resources to gambling researchers, peer reviewers and journal editors.

“One of the problems with systematic reviews is to get it published in a regular gambling journal, you'll get reviewers like me who know the psychology but don't really know... What is there to actually critique in a systematic review? It kind of is what it is. [...] Maybe that's my naive take on something that I do very little in research. Maybe I just don't understand the nuances.” - Aaron, Male researcher based in Canada who studies gambling from a psychological perspective

Introduction

Gambling as a growing public health issue

Gambling is a quickly expanding multinational industry. The global gambling industry grew from \$250 billion USD in 2003 to \$450 billion in 2013 (The Economist, 2014), to a projected \$635 billion by 2022 (Morgan Stanley 2015, as cited in Cassidy, 2020).

Commercial gambling is heavily marketed as a legitimate and harmless leisure activity, but has been called a “dangerous commodity” similar to tobacco and alcohol because of the addiction and other social harms it presents (Markham & Young, 2015). Many people who gamble do not experience harm from gambling; clinically diagnosed problem gambling affects around 2-3% of the global population and has even decreased to 0.6% in Canada (Williams et al., 2020; Williams, Volberg, et al., 2012), but governments and gambling researchers are increasingly approaching gambling as a public health issue as problem gambling is only a small part of the picture.

Firstly, for each person with problem gambling, an average of six other people are harmed by that person’s gambling, usually their closest family and friends (Goodwin et al., 2017). Secondly, although few people gamble at a level that qualifies as clinical problem gambling, many more gamble at what is called a low- or moderate-risk level and still experience harms from that (Browne et al., 2016; Langham et al., 2016).

Thirdly, harms from gambling disproportionately fall onto people in poorer and racialized communities (Abbott et al., 2004), thus the harms from gambling pose an equity issue related to social determinants of health, although the gambling industry’s conflict with public health has been implicated as a “commercial determinant of health” (Ndebele et al., 2020). Finally, there is evidence for a “total consumption model” of gambling harm, meaning there is a strong positive association between the total gambling in a population and harmful gambling in that population (Rossow, 2019; Sulkunen et al., 2019). In light of the gambling industry’s growth over the past 20 years, this means we can expect there has been a concomitant growth in overall gambling-related harms.

Gambling research and the public interest

Governments can use gambling research to inform gambling policies to reduce gambling harm, but there are some limitations to the gambling evidence base. The gambling research funded and conducted in any given country will be shaped by that country’s existing gambling policies. For example, countries that have a history of approaching gambling as a public health issue (e.g., New Zealand) or strong public health research programmes generally (e.g., Nordic countries) produce more public health research on gambling harm prevention, whereas nations like Canada that have a history of viewing gambling as an individual addiction primarily produce research focused on individuals who have gambling problems and how to treat them (Baxter et al., 2019; Nordmyr & Forsman, 2018). Overall, research on the characteristics of individual gamblers is much more common than research on the broader social, cultural, and economic aspects of gambling (Baxter et al., 2019; Hilbrecht & Baxter, 2021).

In addition to this, the gambling industry funds a fair amount of the available gambling research, either by sponsoring research directly or through funding “independent” gambling research organizations. Research funding from the gambling industry has risks of conflict of interest in common with funding from tobacco and alcohol industries (P. J. Adams, 2007; Peter J. Adams, 2016). The 2013 report “Fair Game: Producing gambling research” found that gambling research is behind studies of tobacco, alcohol and other drugs in terms of methods used and dealing with conflicts of interest, and funding

structures tends to focus on “safe” research questions unlikely to lead to further gambling regulation (Cassidy et al., 2013). Two other studies found industry funding had no biasing effect on gambling research, but these studies were themselves funded by the gambling industry and must be approached with caution (Ladouceur et al., 2019; Shaffer et al., 2019).

Finally, for governments and policymakers to utilize gambling research to make evidence-informed decisions, the research must also be summarized and brought to their attention. A common method for summarizing research from multiple research articles on a topic are knowledge syntheses such as systematic reviews and scoping reviews. Unfortunately, the results of the two umbrella reviews evaluating gambling review articles are not promising: one found that primary research articles and systematic reviews on gambling were of low methodological quality in general (McMahon et al., 2019), and the other avoided doing a quality assessment altogether as “most of the studies would have been rated of weak quality and eliminated” (Velasco et al., 2021).

Gambling grey literature

The limitations discussed above largely refer to gambling research and reviews published as primary academic literature, but can be addressed in part by gambling grey literature. Gambling has a sizable and unique body of research published as grey literature: a systematic search of gambling grey literature published in Australia, Canada, New Zealand, the United Kingdom and the United States found that grey literature makes up over 26% of gambling research publications (Baxter et al., 2021b), and a thematic analysis of recent gambling research found that studies published as grey literature tended to focus on broader population health and well-being aspects of gambling, whereas studies published as journal articles tended to focus on the individual characteristics of people who gamble (Baxter et al., 2021a).

Grey literature has other general benefits, such as being more up-to-date and detailed than what is available in the primary literature, as well as providing a “unique global perspective” (Bonato, 2018, p. 28). The latter is especially important in gambling with local research being constrained by the existing policies of the jurisdiction, as context-specific research from different jurisdictions with different gambling policies is more likely to be published as grey literature. For example, the Australian Productivity Commission’s two inquiry reports on gambling (Productivity Commission, 1999, 2010) are highly regarded and well-cited internationally despite their focus on Australian gambling, as they present highly detailed investigative research and analysis that could not be published in an academic journal or book. Somewhat paradoxically, the academic literature aims to make generalized/universal claims, but in doing so loses attention to specific contexts that may inform policies in other contexts. Grey literature such as the Productivity Commission reports achieves this.

Although much of gambling grey research is government-funded and the resulting reports may be consulted by that funder when it is first completed, reuse of those reports is generally lower due to limited distribution of grey literature. One way to ensure that grey research is considered in more gambling policy decisions is to include it in systematic reviews. There are several arguments in favour of this practice.

Firstly, grey literature is a strong contributor to systematic reviews in other health topics (Farrah & Mierzwinski-Urban, 2019; Severn et al., 2017), and the inclusion of grey literature in systematized review articles is consistently considered a best practice and a measure of quality where applicable in prominent review methodologies (Higgins et al., 2021; Kugley et al., 2017; Shea et al., 2017). Due to its significant body of grey literature,

gambling should by default be considered a “where applicable” topic for the inclusion of grey literature in reviews, unless there is an argument for exclusion for a specific gambling topic. Additionally, although grey literature searching is more labour intensive and utilizes more specialized skills, guidelines have been developed for searching health and public health grey literature (Bonato, 2018; Godin et al., 2015) and have been successfully applied to the gambling domain (Baxter et al., 2021b).

Unfortunately, the “grey” of grey literature has been misrepresented as meaning dubious or unpublished in the gambling research community, and excluded from high profile gambling reviews on those grounds (Ladouceur et al., 2017; “Science has a gambling problem,” 2018). On the contrary, grey literature should be included in reviews for the methodological reasons stated above, but also as a matter of principle: Most government-sponsored gambling grey research is funded through governments’ gambling revenues, meaning the research is effectively paid for by gamblers through their gambling losses. Thus, there is an ethical imperative to heed that research in service of reducing gambling-related harms to those gamblers, their families, and their communities.

The present study

I have outlined several reasons why grey literature ought to be included in systematic reviews, as well as some signs that it may not be included to the extent it ought to be. This study presents a citation analysis of recent gambling review articles to answer the following questions:

- To what extent is grey literature cited in gambling review articles?
- To what extent are grey literature sources included in the search strategies of systematized review articles?
 - When included, what grey literature sources are searched?
 - When included, is relevant grey literature found?
- How is grey literature discussed in these articles, if at all?

Methods

All methods presented below were performed solely by the author.

Data source and search strategy

Bibliographic data were collected from the Web of Science (WoS) scholarly database. Of the large multidisciplinary databases Scopus, WoS, and Google Scholar, WoS was chosen because it was the only one to that included “cited references” in bibliographic downloads, and was thus ideal for data processing, although this option is no longer available at the time of publication.

The topic “gambl*” was searched across all WoS indexes, limited to document type “review” and years 2016 to present. The search is restricted to review articles due to our interest in search methodologies, and is in keeping with previous grey literature citation studies (Farrah & Mierzwinski-Urban, 2019; Severn et al., 2017). The initial search was performed December 9th, 2019 and the follow-up search was performed November 9th, 2021.

Eligibility criteria including selection process

Articles were included if they were review articles that investigated gambling or gambling harm (including problem gambling, gambling disorder, and pathological gambling). Reviews that had a scope of gambling and one other condition or activity (e.g., comparing problem gambling and alcohol use disorder) were also included, but articles with a scope

of three or more conditions or a scope of all addictions or addictive behaviours were excluded. Review protocols were included if the full study was not yet published.

Review articles containing the topic keyword “gambl*” were excluded as out of scope if the topic met any of the following criteria:

- Gambling is the method of investigation only. The most common example is the “Iowa Gambling Task”, a laboratory task used to assess decision-making ability that does not resemble any real gambling situation.
- Gambling problems as a side effect of treatment for brain disorders such as Parkinson’s disease
- Gambling in animal models such as rats
- Otherwise unrelated to gambling (e.g., mentions a person with surname “Gamble”, describing a government’s decision as “a gamble”, etc.)

Coding/data collection process/Data items

A Google Form was created and used to input the following data collected from each article. After screening for eligibility, each article was coded for type of review according to Grant and Booth’s (Grant & Booth, 2009) typology of reviews, and divided into systematized reviews (SRs) and non-systematized reviews (NSRs) according to whether or not they employ a systematized search methodology. The most common types of SRs are systematic or scoping reviews, while NSRs are most commonly referred to as literature reviews, narrative reviews, or critical reviews. SRs had additional data collected as described at the end of this section.

To determine whether grey literature is discussed in each review article, the full-text of the article was searched for the following grey literature keywords: “grey”, “gray”, “government”, “document”, “report”, “website”, “conference”, “clinical trial”, “unpublished”, “peer review”, “peer-review”, “dissertation”, “thesis”, and “theses”. The phrase “peer review” is included as it may imply grey literature, based on the misconception that all grey literature is not peer-reviewed. If a keyword match was found, the paragraph was read to confirm that grey literature is referenced, and the text discussing grey literature was copied to the Google Form.

For all articles, the “References” or “Works Cited” section was reviewed by the author to determine how many grey literature works were cited. Greyness was determined in accordance with the Prague Definition of Grey Literature (Schöpfel, 2010), and if greyness could not be determined based on the citation alone, the original work was consulted whenever possible. The number of grey references and total number of references were recorded, and the list of grey citations was copied in the Google Form.

For systematized reviews, additional data was collected. The list of included studies in the Results section was reviewed and the number of grey items and total items included in the synthesis were recorded. The search strategy was also read and all databases potentially containing grey literature were recorded. A source was not recorded if the eligibility criteria excluded all grey literature, or if it was searched in such a way as to exclude grey literature. Data sources could be named databases (e.g., OpenGrey) or generic methods (e.g., snowball searching or contacting subject matter experts). Generic methods were only counted as a grey literature source if either 1) the review explicitly states the purpose was to find grey/unpublished work, or 2) grey literature was present in the included studies. As such, if a review included snowball searching, but the results included no grey literature and the methods did not otherwise acknowledge trying to find grey literature, the snowball searching was not recorded as a grey literature source.

Coding and Analysis

Data from the Google Form were imported to a Microsoft Access database for cleaning and coding. Data for grey literature sources were standardized to a controlled vocabulary and stored in a related table. For the grey literature discourse analysis, systematized articles were coded as to whether or not they included grey literature in the search strategy, and whether they mentioned grey literature in the article text. Where grey literature was discussed, the text was reviewed and the sentiment to grey literature was recorded as one of the following:

- Positive (e.g., “grey literature was included in order to reduce publication bias”)
- Neutral (e.g., “both academic and grey literature sources were searched”)
- Negative (e.g., “grey literature was excluded because it was presumed to be of low quality.
- Conflicted (a combination of positive and negative sentiments)

Summary statistics presented in the following section were created through database queries.

Results

The searches yielded 512 hits for the years 2016-2020. After ineligible articles were excluded, a sample of 174 gambling review articles remained. When coded for review type, 100 reviews had a systematized search strategy, while 74 were not systematized.

A summary of the number of reviews by type and year is shown in Table 1. Overall there is a trend in the number of gambling review articles increasing over time, but this is attributable to a spike in non-systematized review articles published in 2020, likely due to an increase in researchers conducting desk reviews during COVID-19 lockdowns. The number of systematized review articles increased slightly from 2016 to 2017 and remained stable thereafter.

Table 1: Gambling review articles retrieved from the Web of Science database by year, divided into systematized reviews (SRs) and non-systematized reviews (NSRs)

<i>Year</i>	<i>SRs</i>	<i>NSRs</i>	<i>All reviews</i>
<i>2016</i>	15	8	23
<i>2017</i>	22	14	36
<i>2018</i>	22	8	30
<i>2019</i>	20	12	32
<i>2020</i>	21	32	53
<i>All years</i>	<i>100</i>	<i>74</i>	<i>174</i>

Do systematized review search strategies include grey literature?

For the sample of 100 SRs, Table 2 shows how often grey literature sources are included in the search strategy. Just over half of the SRs included grey literature (54 of 100). There is no significant trend in grey literature inclusion increasing or decreasing over time, neither in raw numbers of SRs including grey literature or as a proportion of all SRs.

Table 2: Systematized review articles on gambling retrieved from the Web of Science database (n=100), summarized by year and whether grey literature is included in the search strategy.

<i>Year</i>	<i>Includes grey literature</i>	<i>Excludes grey literature</i>	<i>Proportion including grey literature</i>
2016	9	6	0.6
2017	11	11	0.5
2018	14	8	0.64
2019	10	10	0.5
2020	10	11	0.48
<i>All years</i>	<i>54</i>	<i>46</i>	<i>0.54</i>

What grey literature sources are searched?

Across the 54 reviews that included grey sources, 90 different sources were searched. The top 21 sources, which were employed in 4 or more reviews, are listed in Table 3. Of the other 69 sources, 44 were employed in only one review, 18 in two reviews, and 6 in three reviews. The full list of sources searched will be available in the published dataset.

Table 3: The most common sources of grey literature in gambling systematized reviews that search grey literature, 2016-2020 (n=54)

<i>Source name</i>	<i># of times searched</i>	<i>Source name (cont'd)</i>	<i># of times searched</i>
Google Scholar	17	GambLib.org	6
Google	12	McGill International Centre for Youth Gambling Problems	5
GREO Evidence Centre	11	[generic] Contact with experts	5
ProQuest Dissertations & Theses	11	Gambling Commission (Great Britain)	5
[generic] Snowball searching	10	Gordon Moody Association	5
Responsible Gambling Council (Ontario)	9	National Problem Gambling Clinic (UK)	5
Open Grey	8	GamCare	5
GambleAware InfoHub	8	WHO International Clinical Trials Registry Platform	4
Alberta Gambling Research Institute (AGRI) Repository	6	Advisory Board for Safer Gambling	4
ClinicalTrials.gov	6	Gamblers Anonymous	4
Victorian Responsible Gambling Foundation	6		

The two most popular grey literature sources were Google and Google Scholar, while several gambling-specific organization websites were also commonly used. A few general grey literature sources were popular, including ProQuest for theses and dissertations, two registries for unpublished clinical trials, OpenGrey for general grey literature, and snowball searching and contacting subject matter experts to find additional (including grey) material.

Figure 1 summarizes how many grey literature sources were searched in each article. Of the 54 systematized review articles that searched grey literature, nearly 60% (32 of 54) only searched one or two grey literature sources, which were most commonly a simple query of Google or Google Scholar or a snowball search of the references in included studies. This is in addition to the remaining 46 systematized reviews that did not search grey literature at all.

However, there were also some thorough grey literature search strategies employed. Seven articles searched over 10 grey literature sources, while three searched over 20. The top five articles by number of grey sources search each represent high quality review methods generally and grey literature search methods specifically, and provide useful models for gambling grey literature inclusion (Beynon et al., 2020a, 2020b; Bramley et al., 2018; Wardle et al., 2019; Yakovenko & Hodgins, 2018).

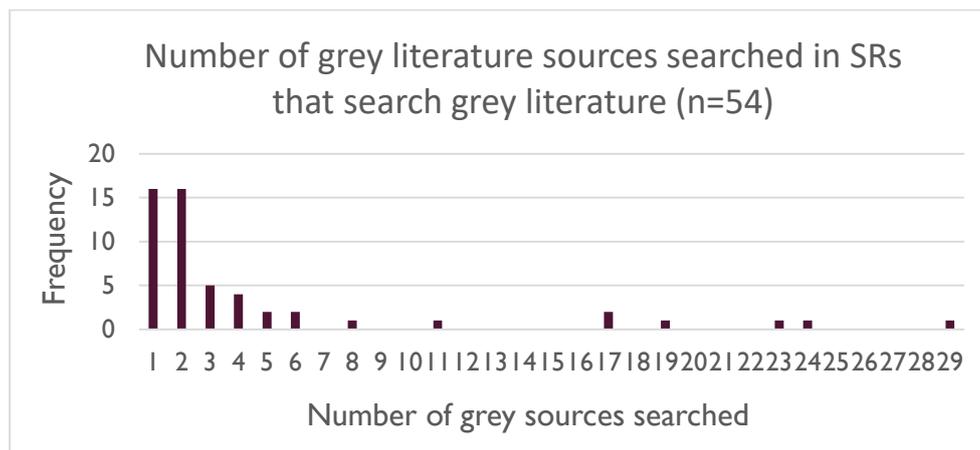


Figure 1: Distribution of number of grey literature sources searched in gambling systematized reviews that include grey literature, 2016-2020 (n=54)

To what extent is grey literature cited?

Of the 54 SRs that included grey literature in the search strategy, nearly three quarters (40 of 54) ultimately included one or more grey publication in the included studies. This is encouraging considering that 32 (nearly 60%) of these reviews only searched one or two grey literature sources, so the proportion of gambling review topics for which relevant grey literature exists is likely higher than what is reported here.

When looking at individual studies included in systematic review syntheses, a total of 3961 citations were made. Of these, 353 were citations of grey literature, or approximately 9% of included studies. This increases to 12% when only SRs that included grey literature are considered (352 of 2,385 included studies).

When considering the full references list of gambling reviews, grey literature represents 9% of all works cited (1,257 of 13,728 citations; see Table 4). When broken down by type of review, the proportion ranges from 14% in SRs that include grey literature, to 7% both for SRs that do not include grey literature as well as NSRs. This shows that for all types of

reviews, including systematized reviews that exclude grey literature from the search strategy, there is a certain baseline of grey literature citation in the introduction, methods, and/or discussion sections. However, the total proportion of grey literature cited is much lower than the 26-47% that has been found in systematic reviews of other health domains (Farrah & Mierzwinski-Urban, 2019; Severn et al., 2017), and is also lower than the 18% in a sample of research publications on education (Češarek & Merčun Kariž, 2019).

Table 4: Summary of number of grey literature and total works cited in a sample of gambling review articles (N=174), according to type of review.

Type of article	Grey citations	Total citations	% grey citations
SRs that include grey literature (n=54)	569	4032	14.1%
SRs that exclude grey literature (n=46)	259	3672	7.1%
NSRs (n=74)	429	6024	7.1%
All articles (N=174)	1,257	13,728	9.2%

How is grey literature discussed in gambling review articles?

Within this sample, grey literature is mentioned or discussed in 76 of 100 systematized reviews (76%), and only 6 of 74 non-systematized reviews (8%). It is not expected that grey literature be mentioned or discussed in non-systematized reviews since they normally do not report search methods. Thus, this section will focus on systematized reviews only.

Table 5 summarizes how grey literature is referred to and discussed in gambling systematized reviews. When grey literature is included it is usually acknowledged neutrally, but over one quarter of these articles mention some positive aspect of including grey literature, such as providing additional evidence or reducing publication bias. One review included grey literature but in the study quality assessment scheme assigned grey literature a lower score (Kotter et al., 2019), so this is classified as a negative sentiment. Interestingly, seven reviews included grey literature through a general source such as Google Scholar but did not acknowledge its greyness in any way.

In reviews that did not include grey literature, there were many more articles that did not mention grey literature at all (17 of 46, or 37%). Another 19 articles discussed grey literature neutrally by acknowledging that grey literature was excluded or that only peer-reviewed literature was included, without any reasoning or elaboration. Only one review, the umbrella review by McMahon et al. (2019), provided a neutral methodological justification for excluding grey literature. Only five articles, including one by the present author, were positive about grey literature by acknowledging its exclusion as a limitation of the study, while two were negative in that grey literature was excluded because it was assumed to be of lower quality than academic journal articles. Finally, three reviews were classified as “conflicted” in their sentiment towards grey literature, because they state the exclusion of grey literature was a limitation of the study while simultaneously calling in to question the quality or suitability of grey literature.

Table 5: Summary of sentiments towards grey literature in gambling systematized review articles (n=100)

Review type	Sentiment	n	Description or Examples
Grey literature included (n=54)	Positive	14	Grey literature as source of evidence beyond the published literature (n=5) Grey literature addresses publication bias (n=5) Grey literature as a significant source of gambling literature generally (n=1)
	Neutral	32	Specific grey literature sources included and acknowledged as such Sources that include both primary and grey literature (e.g., Google Scholar), acknowledged as a source of grey literature.
	Negative	1	Grey literature included in review but assigned a lower score in the study quality assessment scheme
	N/A	7	Sources that include grey literature (e.g., Google Scholar) are included but not acknowledged as sources of grey literature, and grey literature is not subsequently excluded
Grey literature excluded (n=46)	Positive	5	Grey literature not included but this is noted as a limitation of the study.
	Neutral	19	Grey sources or “non-peer reviewed” sources excluded, with no reasoning given <i>“Only peer-reviewed articles published in academic journals [are included]”</i> . Grey literature excluded by omission One article, an umbrella review, appealed to previous umbrella reviews only reviewing primary literature (McMahon et al., 2019)
	Negative	2	Appeal to the assumption that journal articles are more rigorous than grey literature as grounds for exclusion. <i>“To be included as an output to be evaluated, the published paper had to have: [...] (iv) been subjected to peer-review. It was assumed that those studies that had undergone peer-review would be more scientifically rigorous than anything in the “grey” literature”</i> (Harris & Griffiths, 2018)
	Conflicted	3	Exclusion of grey literature discussed in limitations section, but the quality or suitability of grey literature is also questioned <i>“The current review was also limited by its inclusion of only peer-reviewed work. It remains possible that books, dissertations, or grey literature could provide more detail [on the topic] given the multi-disciplinary nature of the field. However, these forms of media were excluded to ensure a consistent level of quality throughout the review.”</i> (Barton et al., 2017)
	N/A	17	Grey literature not included nor discussed

Discussion

In previous work we have shown that grey literature constitutes a sizable and unique portion of research produced about gambling (Baxter et al., 2021a). This study extends that previous work by analyzing how that grey literature evidence base is incorporated into knowledge syntheses (i.e., review articles) about gambling. Grey literature was found to play an important role in knowledge syntheses but is perhaps underutilized.

Although grey literature was searched in just over half of systematized reviews on gambling, most of these grey literature search strategies were not comprehensive. Furthermore, the proportion of reviews including grey literature did not increase over the five-year period investigated.

Grey literature represented 9% of the “included” studies in systematized reviews, and this increases to 12.5% when only considering those that included grey sources at all. Grey literature also represented 9% of all works cited across all types of reviews. These figures are lower than results from similar grey literature analyses in other health-related

domains. A study sampling review articles employing Godin et al.'s (2015) grey literature search methodology found that grey literature made up 23% of documents cited (Severn et al., 2017). This figure jumps to 47% for horizon scanning reports on non-drug health technologies (Farrah & Mierzwinski-Urban, 2019), although it is reduced to 33% when excluding manufacturer information, which would have limited applicability in reviews on gambling harm-related topics .

The lower grey literature citation rate in gambling could be because there is simply less relevant grey literature available, but it is likely at least in part due to less thorough methods. In Severn et al.'s (2017) sample a grey literature search methodology and checklist was always used, whereas in the gambling sample grey literature searches were only employed in just over half of systematized reviews, and were usually not thorough. Despite these limitations, at least one relevant grey document was found in nearly three quarters of gambling systematized reviews that searched for it. Taken together this suggests that more relevant grey literature may be found and included in gambling review articles if a grey literature search methodology was applied more often.

The discourse analysis revealed some further concerning results. I have argued that grey literature is relevant to many aspects of gambling harm and thus for most gambling review articles a grey literature search is required for the search strategy to be considered comprehensive. This study found that of 46 systematized reviews that did not include grey literature, 17 did not mention grey literature at all, and of the other 29 only one was judged to give a sufficient methodological justification for its exclusion. The three studies with a "conflicted" position on grey literature are interesting as they demonstrate a tension between understanding of grey literature's value and reservations about its quality.

Thus, it is not only the case that grey literature ought to be included in gambling review articles more often, but also when it is excluded it should be discussed as a limitation of the study, or some methodological reason should be given, even if the reason is that a grey literature search was beyond the scope of the article or the means of the authors.

As this study has focused on academic journal articles, the opportunities for improvement lie with the study authors, journal editors, and peer reviewers. For this reason, attention should be paid to the journals in which gambling research is published. A mapping review of 2,266 articles investigating antecedents to harmful gambling found that gambling research is overwhelmingly published in gambling-specific or addiction journals (Hilbrecht & Baxter, 2021). Similarly, of the 23 "generally of low quality" review articles in the two umbrella reviews on gambling (McMahon et al., 2019; Velasco et al., 2021), all but two were published in a gambling or addiction journal (the exceptions being Williams, West, et al., (2012) and Livingstone et al., (2014)). Although detailed analysis is not presented here, 29% of the systematized reviews in the current study sample were published in gambling journals, while another 31% were published in addiction or behavioural addiction journals, and the rest distributed across psychology, psychiatry, neuroscience, medicine, public health, and social sciences journals. Thus, it is possible that grey literature methods are not as well understood utilized in gambling and behavioural addictions research communities as they are in the broader health research community.

Opportunities for improvement

For researchers writing gambling review articles, it is of course recommended to include grey literature sources in the search strategy. Godin et al.'s (2015) methodology serves as a strong basis for grey literature searches treating gambling as a public health issue, and if a thorough grey literature search is not possible, our previous application of the

methodology (Baxter et al., 2021b) has shown that Greo's International Gambling Research Evidence Centre (Greo, n.d.) and the Alberta Gambling Research Institute Repository (Alberta Gambling Research Institute, n.d.) have strong international coverage and are functional and up-to-date at the time of writing, and are recommended as baseline gambling grey literature sources.

Review methods can also be improved generally through the assistance of an information specialist. A study of general medicine systematic reviews found that those with a librarian coauthor had the highest quality search strategies, and those with a librarian listed in the "Acknowledgements" section were still significantly higher quality than those with no librarian involvement (Rethlefsen et al., 2015). Thus, researchers are encouraged to take advantage of any systematic review services that may be available through their organization's library.

As the value of peer-review has been raised in many gambling systematized reviews excluding grey literature, another method to improve one's review is to seek peer-review from those who are experienced with review methods. Systematic reviews can be preregistered in PROSPERO and some journals will accept review protocols for peer review and publication (e.g., *Systematic Reviews*, see Beynon et al., 2020a, 2020b). Although it is often not feasible to publish a protocol, proactive peer-review for review methods is available through PRESS (Peer Review of Electronic Search Strategies, McGowan et al., 2016). Proactive peer review was encouraged in the original PRISMA statement (Moher et al., 2009), but was expanded to its own reporting item in the 2021 PRISMA-S extension in acknowledgement of its value in reducing errors and potentially increasing the number of included relevant studies (Rethlefsen et al., 2021). Although one study found that peer review only added additional relevant studies in 4% of a sample of rapid reviews (Spry & Mierzwinski-Urban, 2018), the current study suggests that peer review would bring greater benefits to the "generally low quality" systematic reviews on gambling, especially if it resulted in the addition of grey literature sources.

The article in the present sample with the most thorough grey literature search strategy, Beynon et al. (2020a), is an exemplar for all the above recommendations and is available Open Access. It is a protocol published in the journal *Systematic Reviews*, meaning the methods were preregistered and received external peer-review of the search strategy before any searches were performed. It acknowledges that the search strategy was developed by a senior information scientist and internally peer-reviewed by a second information specialist. Not surprisingly, this review protocol has thorough grey literature searches that include the baseline gambling websites recommended above. It is recommended that this protocol be referenced by anyone aiming to perform a high-quality systematic search of gambling literature.

These improvements can likewise be encouraged by peer reviewers and editors of gambling and addiction journals. As peer review normally happens after a review article has been completed, it may not be feasible for a peer reviewer to require revisions that include a comprehensive grey literature search. If there is no grey literature search in a submitted systematic review, peer reviewers could ask for a simple search of the "baseline" gambling databases listed above, or at minimum should require that the exclusion of grey literature be noted as a limitation. Peer reviewers not experienced with systematized review methods should refer to the PRISMA-S reporting guidelines checklist while reviewing (Rethlefsen et al., 2021).

Editors of gambling and addiction journals could provide these guidelines and resources to subject-matter-expert peer reviewers, but can also seek to recruit information

specialists as methodology-expert peer reviewers for systematic review articles. Librarians are an underutilized source of peer-review expertise: A recent survey found that less than a quarter of health librarians had been approached to do peer review, but their reviews almost always result in revisions or rejection based on the search methods, while 95% of those who hadn't been approached said they would (54%) or might (41%) peer review (Grossetta Nardini et al., 2019).

Gambling and addiction journal editors are also encouraged to acknowledge the value of grey literature in editorial policy documents, for systematic reviews as well as an important source of evidence for narrative reviews and research article introductions. The new journal *Critical Gambling Studies* explicitly and positively acknowledges grey literature in the article assessment rubric provided to authors and peer reviewers ("Submissions," n.d.). Editors of other journals may use this as a model and are also encouraged to escalate grey literature guidelines and standards to their peers and communities of practice, such as the International Society of Addiction Journal Editors (ISAJE).

Limitations and further research

This study has noteworthy limitations which present opportunities to build on or further analyze the existing dataset. Firstly, the scope of the study was limited to a five-year period and only one multidisciplinary article database, Web of Science, which was selected because it provided the most relevant bibliographic data. Future analyses could extend the time period and searching more databases. Scopus is recommended as the next scholarly database to include as it has broad and complementary multidisciplinary coverage and the necessary search functions.

It would be remiss and hypocritical to not acknowledge that this review does not itself include grey literature. Although the focus on journal articles provides specific insight into peer review and publishing practices of scholarly journals, there are also many gambling review studies that are published as grey literature reports that are better positioned to inform and influence policy decisions. Thus, it would be valuable to also investigate to what extent grey gambling reviews themselves utilize grey literature and what the opportunities for improvement may be there. Again, the Greo and AGRI gambling grey literature databases provide excellent baseline sources for this investigation.

This study did not differentiate between disciplinary approaches to gambling research other than the preliminary analysis of journals. Further analysis by discipline will be particularly revealing for biological approaches to gambling, such as neuroscientific and animal model studies, which are often published outside the popular journals for gambling (Hilbrecht & Baxter, 2021). In the current study, grey literature inclusion and citation was low in both systematized and non-systematized reviews on these topics. This field of gambling research merits further investigation as it may also have a different yet also underutilized body of grey literature (for example, mainly preprints, conference proceedings, or clinical trials rather than government and institute reports).

A broad variety of grey literature sources were consulted across the reviews in this study, but these were not investigated in detail. An analysis comparing the number and types of grey literature sources searched to the number of relevant grey resources found would provide valuable insight into which grey literature sources are most useful for gambling, and the number of relevant sources that constitutes a sufficiently comprehensive gambling grey literature search.

Finally, future analysis of the dataset could catalogue the individual grey literature items cited in this sample and the number of times each item is cited. This would provide an informative impact assessment as to what grey gambling publications are cited the most often. These data are not readily available for bodies of grey literature as they are for journal articles and can provide insight into the attributes of the most influential and potentially highest quality gambling grey literature.

Acknowledgements

This study was supported by an Alberta Gambling Research Institute Graduate Scholarship to the author.

The pseudonymous epigraph was collected and published as part of the project “Revealing Gambling’s Grey Literature”, approved by the University of Alberta Research Ethics Board (Pro #00107486).

References

- Abbott, M. W., Volberg, R., Bellringer, M. E., & Reith, G. (2004). *A review of research on aspects of problem gambling*. London: Retrieved from https://niphmhr.aut.ac.nz/__data/assets/pdf_file/0004/7537/summary_of_auckland_report.pdf
- Adams, P. J. (2007). Assessing whether to receive funding support from tobacco, alcohol, gambling and other dangerous consumption industries. *Addiction*, *102*(7), 1027-1033. <https://doi.org/10.1111/j.1360-0443.2007.01829.x>
- Adams, P. J. (2016). *Moral Jeopardy: Risks of Accepting Money from the Alcohol, Tobacco and Gambling Industries*. Cambridge: Cambridge University Press.
- Alberta Gambling Research Institute. (n.d.). Alberta Gambling Research Institute Repository. Retrieved from <https://prism.ucalgary.ca/handle/1880/100000>
- Barton, K. R., Yazdani, A., Ayer, N., Kalvapalle, S., Brown, S., Stapleton, J., . . . Harrigan, K. A. (2017). The Effect of Losses Disguised as Wins and Near Misses in Electronic Gaming Machines: A Systematic Review. *Journal of Gambling Studies*, *33*(4), 1241-1260. <https://doi.org/10.1007/s10899-017-9688-0>
- Baxter, D. G., Hilbrecht, M., & Wheaton, C. T. J. (2019). A mapping review of research on gambling harm in three regulatory environments. *Harm Reduction Journal*, *16*, 19. <https://doi.org/10.1186/s12954-018-0265-3>
- Baxter, D. G., Nicoll, F., & Akcayir, M. (2021a). Grey Literature is a Necessary Facet in a Critical Approach to Gambling Research. *The Grey Journal*, *17*(3).
- Baxter, D. G., Nicoll, F., & Akcayir, M. (2021b, Apr. 27-29). *A protocol for identifying gambling grey literature for systematic and scoping reviews*. Paper presented at the Alberta Gambling Research Institute Annual Conference 2021: Gambling in Canada: Current Research & Future Directions, Online Conference.
- Beynon, C., Pearce-Smith, N., & Clark, R. (2020a). Harms associated with gambling: abbreviated systematic review protocol. *Systematic Reviews*, *9*(1). <https://doi.org/10.1186/s13643-020-01397-4>
- Beynon, C., Pearce-Smith, N., & Clark, R. (2020b). Risk factors for gambling and problem gambling: a protocol for a rapid umbrella review of systematic reviews and meta-analyses. *Systematic Reviews*, *9*(1). <https://doi.org/10.1186/s13643-020-01455-x>
- Bonato, S. (2018). *Searching the Grey Literature: A Handbook for Searching Reports, Working Papers, and Other Unpublished Research*. New York: Rowman and Littlefield.
- Bramley, S., Norrie, C., & Manthorpe, J. (2018). Gambling-related harms and homelessness: findings from a scoping review. *Housing Care and Support*, *21*(1), 26-39. <https://doi.org/10.1108/hcs-02-2018-0003>
- Browne, M., Langham, E., Rawat, V., Greer, N., Li, E., Rose, J., . . . Best, T. (2016). *Assessing gambling-related harm in Victoria: A public health perspective*. Melbourne: Retrieved from https://www.responsiblegambling.vic.gov.au/__data/assets/pdf_file/0007/28465/Browne_assessing_gambling-related_harm_in_Vic_Apr_2016-REPLACEMENT2.pdf
- Cassidy, R. (2020). *Vicious Games: Capitalism and Gambling*. London: Pluto Press.

- Cassidy, R., Loussouarn, C., & Pisac, A. (2013). *Fair Game: Producing Gambling Research*. London: Retrieved from https://prism.ucalgary.ca/bitstream/handle/1880/50242/Fair_Game_Web_Final.pdf?sequence=1&isAllowed=y
- Češarek, A., & Merčun Kariž, T. (2019). *Grey literature cited in scientific articles: case study at the Faculty of Education, University of Ljubljana, Slovenia*. Paper presented at the 21st International Conference on Grey Literature, Hanover, Germany.
- Farrah, K., & Mierzwinski-Urban, M. (2019). Almost half of references in reports on new and emerging nondrug health technologies are grey literature. *Journal of the Medical Library Association*, 107(1), 6. <https://doi.org/10.5195/jmla.2019.539>
- Godin, K., Stapleton, J., Kirkpatrick, S. I., Hanning, R. M., & Leatherdale, S. T. (2015). Applying systematic review search methods to the grey literature: A case study examining guidelines for school-based breakfast programs in Canada. *Systematic Reviews*, 4(1), 138. <https://doi.org/10.1186/s13643-015-0125-0>
- Goodwin, B. C., Browne, M., Rockloff, M., & Rose, J. (2017). A typical problem gambler affects six others. *International Gambling Studies*, 17(2), 276-289. <https://doi.org/10.1080/14459795.2017.1331252>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91-108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Greo. (n.d.). Evidence Centre. Retrieved from <https://www.greo.ca/en/greo-resource/evidence-centre.aspx>
- Grossetta Nardini, H. K., Batten, J., Funaro, M. C., Garcia-Milian, R., Nyhan, K., Spak, J. M., . . . Glover, J. G. (2019). Librarians as methodological peer reviewers for systematic reviews: results of an online survey. *Research Integrity and Peer Review*, 4(1), 23. <https://doi.org/10.1186/s41073-019-0083-5>
- Harris, A., & Griffiths, M. D. (2018). The Impact of Speed of Play in Gambling on Psychological and Behavioural Factors: A Critical Review. *Journal of Gambling Studies*, 34(2), 393-412. <https://doi.org/10.1007/s10899-017-9701-7>
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2021). *Cochrane Handbook for Systematic Reviews of Interventions* [version 6.2 (updated February 2021)]. Retrieved from www.training.cochrane.org/handbook
- Hilbrecht, M., & Baxter, D. G. (2021, Apr. 27-29). *Exploring Trends in Canadian Gambling Research Publications: Alignment with the Conceptual Framework of Harmful Gambling over a 12-Year Period*. Paper presented at the Alberta Gambling Research Institute Annual Conference 2021: Gambling in Canada: Current Research & Future Directions, Online Conference.
- Kotter, R., Kraplin, A., Pittig, A., & Buhringer, G. (2019). A Systematic Review of Land-Based Self-Exclusion Programs: Demographics, Gambling Behavior, Gambling Problems, Mental Symptoms, and Mental Health. *Journal of Gambling Studies*, 35(2), 367-394. <https://doi.org/10.1007/s10899-018-9777-8>
- Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A.-M. K., Hammerstrøm, K., & Sathe, N. (2017). Searching for studies: a guide to information retrieval for Campbell systematic reviews. *Campbell Systematic Reviews*, 13(1), 1-73. <https://doi.org/10.4073/cmg.2016.1>
- Ladouceur, R., Shaffer, P., Blaszczynski, A., & Shaffer, H. J. (2017). Responsible gambling: A synthesis of the empirical evidence. *Addiction Research & Theory*, 25(3), 225-235. <https://doi.org/10.1080/16066359.2016.1245294>
- Ladouceur, R., Shaffer, P., Blaszczynski, A., & Shaffer, H. J. (2019). Responsible gambling research and industry funding biases. *Journal of Gambling Studies*, 35(2), 725-730. <https://doi.org/10.1007/s10899-018-9792-9>
- Langham, E., Thorne, H., Browne, M., Donaldson, P., Rose, J., & Rockloff, M. (2016). Understanding gambling related harm: a proposed definition, conceptual framework, and taxonomy of harms. *Bmc Public Health*, 16(1), 80. <https://doi.org/10.1186/s12889-016-2747-0>
- Livingstone, C., Rintoul, A., & Francis, L. (2014). What is the evidence for harm minimisation measures in gambling venues? *Evidence Base*(2). Retrieved from <https://search.informit.org/doi/pdf/10.3316/informit.209415982597362>
- Markham, F., & Young, M. (2015). "Big Gambling": The rise of the global industry-state gambling complex. *Addiction Research & Theory*, 23(1), 1-4. <https://doi.org/10.3109/16066359.2014.929118>
- McGowan, J., Sampson, M., Salzwedel, D. M., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *Journal of Clinical Epidemiology*, 75, 40-46. <https://doi.org/10.1016/j.jclinepi.2016.01.021>

- McMahon, N., Thomson, K., Kaner, E., & Bimbra, C. (2019). Effects of prevention and harm reduction interventions on gambling behaviours and gambling related harm: An umbrella review. *Addictive Behaviors, 90*, 380-388. <https://doi.org/10.1016/j.addbeh.2018.11.048>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The Prisma Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine, 6*(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Ndebele, P., Shaikh, H., Paichadze, N., Bari, I., Michaels, D., Burgoa, C. S., & Hyder, A. A. (2020). Commercial determinants of health: an ethical exploration. *International Journal of Public Health, 65*(7), 1123-1132. <https://doi.org/10.1007/s00038-020-01427-x>
- Nordmyr, J., & Forsman, A. K. (2018). A systematic mapping of nordic gambling research 2000-2015: current status and suggested future directions. *Addiction Research & Theory, 26*(5), 339-348. <https://doi.org/10.1080/16066359.2018.1426753>
- Productivity Commission. (1999). *Australia's Gambling Industries*. Canberra, AU: Retrieved from <https://www.pc.gov.au/inquiries/completed/gambling/report>
- Productivity Commission. (2010). *Gambling*. Canberra, AU: Retrieved from <https://www.pc.gov.au/inquiries/completed/gambling-2010/report>
- Rethlefsen, M. L., Farrell, A. M., Osterhaus Trzasko, L. C., & Brigham, T. J. (2015). Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *Journal of Clinical Epidemiology, 68*(6), 617-626. <https://doi.org/10.1016/j.jclinepi.2014.11.025>
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., . . . Prisma-S. Group. (2021). PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. *Systematic Reviews, 10*(1), 39. <https://doi.org/10.1186/s13643-020-01542-z>
- Rossow, I. (2019). The total consumption model applied to gambling: Empirical validity and implications for gambling policy. *Nordic Studies on Alcohol and Drugs, 36*(2), 66-76. <https://doi.org/10.1177/1455072518794016>
- Schöpfel, J. (2010). *Towards a Prague Definition of Grey Literature*. Paper presented at the Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues., Prague, Czech Republic. https://archivesic.ccsd.cnrs.fr/sic_00581570/document
- Science has a gambling problem. (2018). *Nature, 553*(379).
- Severn, M., Farrah, K., Ford, C., McCormack, S., Aleksandra, G., & Kaunelis, D. (2017). *Fishing for grey literature: What are we catching in CADTH's rapid response service?*. Paper presented at the Canadian Health Libraries Association/Association des bibliothèques de la santé du Canada (CHLA/ABSC) 2017 Annual Conference: Northern Illumination/Lumières du Nord, Edmonton AB, Canada.
- Shaffer, P. M., Ladouceur, R., Williams, P. M., Wiley, R. C., Blaszczynski, A., & Shaffer, H. J. (2019). Gambling research and funding biases. *Journal of Gambling Studies, 35*(3), 875-886. <https://doi.org/10.1007/s10899-019-09875-8>
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., . . . Henry, D. A. (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ, 358*, j4008. <https://doi.org/10.1136/bmj.j4008>
- Spry, C., & Mierzwinski-Urban, M. (2018). The impact of the peer review of literature search strategies in support of rapid review reports. *Research Synthesis Methods, 9*(4), 521-526. <https://doi.org/10.1002/jrsm.1330>
- Submissions. (n.d.). Retrieved from <https://criticalgamblingstudies.com/index.php/cgs/about/submissions>
- Sulkunen, P., Babor, T. F., Cisneros Ornberg, J., Egerer, M., Hellman, M., Livingstone, C., . . . Rossow, I. (2019). *Setting Limits: Gambling, Science and Public Policy*. Oxford: Oxford University Press.
- The Economist. (2014). Daily chart: The house wins. Retrieved from <http://www.economist.com/blogs/graphicdetail/2014/02/daily-chart-0>
- Velasco, V., Scattola, P., Gavazzeni, L., Marchesi, L., Nita, I. E., & Giudici, G. (2021). Prevention and Harm Reduction Interventions for Adult Gambling at the Local Level: An Umbrella Review of Empirical Evidence. *International Journal of Environmental Research and Public Health, 18*(18), 9484. Retrieved from <https://www.mdpi.com/1660-4601/18/18/9484>
- Wardle, H., Bramley, S., Norrie, C., & Manthorpe, J. (2019). What do we know about gambling-related harm affecting migrants and migrant communities? A rapid review. *Addictive Behaviors, 93*, 180-193. <https://doi.org/10.1016/j.addbeh.2019.01.017>

- Williams, R. J., Leonard, C. A., Belanger, Y. D., Christensen, D. R., el-Guebaly, N., Hodgins, D. C., . . . Stevens, R. M. G. (2020). Gambling and Problem Gambling in Canada in 2018: Prevalence and Changes Since 2002. *The Canadian Journal of Psychiatry*, 66(5), 485-494. <https://doi.org/10.1177/0706743720980080>
- Williams, R. J., Volberg, R., & Stevens, R. (2012). *The Population Prevalence of Problem Gambling: Methodological Influences, Standardized Rates, Jurisdictional Differences, and Worldwide Trends*. Guelph, CA: Retrieved from <https://www.greo.ca/Modules/EvidenceCentre/Details/population-prevalence-problem-gambling-methodological-influences-standardized-rates-0>
- Williams, R. J., West, B. L., & Simpson, R. (2012). *Prevention of Problem Gambling: A Comprehensive Review of the Evidence and Identified Best Practices*. Guelph, CA: Retrieved from <https://opus.uleth.ca/handle/10133/3121>
- Yakovenko, I., & Hodgins, D. C. (2018). A scoping review of co-morbidity in individuals with disordered gambling. *International Gambling Studies*, 18(1), 143-172. <https://doi.org/10.1080/14459795.2017.1364400>

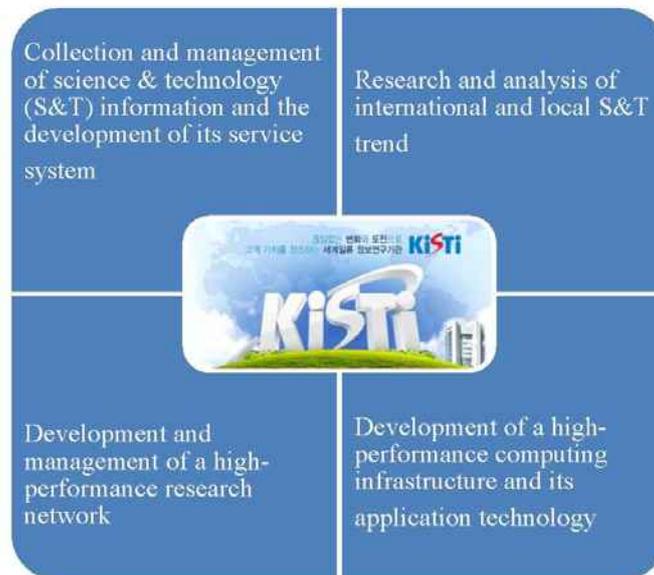
Korea Institute of Science and Technology Information (KISTI)

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

* Vision

World-class information research institute creating values for customers

* Main functions



* Management and service of Korean R&D reports

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

* Contact information

KISTI email address: hcpark@kisti.re.kr

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

Digital Grey Soviet Science

Yuliya B. Balashova, Saint Petersburg State University, Russia

Abstract

The article examines Soviet science representation on the leading Russian scientific and educational network platforms. Thematic dominants in the appropriate GL materials are considered.

Key words: *Soviet science, popular science digital, grey literature.*

Relevance

In the global pandemic context, public science becomes primary important, being represented, among other things, on the various network platforms. The lack of this kind relevant scientifically grounded information was clearly manifested all over the world and totally negative contributed to the COVID-dissidence moods. Russian pandemic situation specificity correlates with the fact that Russia was the first country created the vaccine, where everyone could freely and quickly have been vaccinated. However, such a unique achievement encounters a certain opposition of the population, caused mainly the distancing from the social institutions of society. Crisis of confidence in official science is especially typical for the former USSR countries (with the different political regimes). COVID-19 topic is extremely politicized, still little implemented in the field of public science communication. To contradict this situation, in Russia, the last 2021 year, was announced as The Year of Science and Technology.

For all advantages, digital environment has tendency to destroy value hierarchy, authority as such, and the opinion of the honored leaders. On this background, occurs institution of trust and authority (key for the Russian culture) destruction, the general tendency towards the variability of historical memory escalates (for example, “history on demand” phenomenon).

In the aspect of continuation, it must be stated that the modern people world view (for instance, in the U.S.) paradoxically remains in some way Soviet. Recently, sociologists recorded the increase sympathy for the Soviet project among Russians. It's noteworthy that the USSR was mainly constituted by the cult of science. The Soviet social and scientific projects, leadership position in the world make the USSR attractive today. However, over the 30 years that have passed since the USSR had collapsed, some groups of people, unfortunately, have lost confidence into fundamental scientific knowledge.

Russian vaccines effectiveness (first of all – “Sputnik V”) is provided by the rich Soviet virology legacy. So, into this context, various digital publications about Soviet science appeared, including successful Soviet experience in combating pandemics.

Organization of Research

The goal is to reveal Soviet science modern digital media representation. The primary task is to compare the main Soviet popular science thematic areas with their manifestation on the most popular modern scientific and educational network platforms due GL materials. The GL materials genre composition includes digital articles, podcasts, lectures, essays. Different science communicators has been involved into such popular science projects, partly (on some platforms – mostly) – scientists themselves.

The research method is associated with problematic and thematic analysis of the Russian popular science digital media, in which Soviet science is being updated. The main areas of semiotics (considering semiotic project importance in the USSR) in relation to the Soviet

science media representation are observed. Syntactic aspect of the media representation includes Soviet science organization; semantic aspect relates to the Soviet science symbolic potential; pragmatic aspect connects to the socio-political agenda.

Popular science and educational network platforms: *PostScience (PostNauka)*, *Arzamas*, *Diletant.media* – created in Russia as independent media, not based on the Western franchise, were selected as the sources.

Historical digital archives importance repeatedly indicated in the modern scientific literature (4).

Discussion

As we know, science claims to create a matrix of the 20th century human consciousness. The Soviet regime considered one of the priority tasks to develop people's scientific world outlook. The Soviet scientific project turned out to be extremely successful, among other things, due to the comprehensive and deep science mediatization. Since the USSR formation, the mission was to raise Soviet people mind to the scientific. This task was successfully solved; in the late Soviet period of the 1970s and 1980s, the level of scientific thinking among the population was extremely high. An integral system of science mediatization has developed in the USSR. All possible communication channels were included into the process of the Soviet science mediatization (5). First of all, there were such institutions of society, as education, enlightenment, and media. Soviet scientists were obliged to engage in scientific popularization, this was considered their public duty. Academicians have written articles for the press, provided public lectures. A number of famous scientists were at the same time the most prominent knowledge popularizers (like the physicist *Sergei Kapitsa*). Science became a part of mass culture and everyday life of the Soviet people.

Syntactic aspect of the media representation.

In the USSR, an effective science policy has been formed. Thanks to this strategy, subcultural niches appeared, represented "state within a state", alternative to the dominant ideology. The academy could contradict the authorities. The unique environment of research institutes and "science cities" was created. The modern actionist and scandalous art project "Dau", named after the Russian physicist *Leo Landau*, successfully held in Europe, also reconstructed this unique matrix with its own methods. So, to form his own school of physics, Landau invented the author's exam, the "theoretical minimum". For example, he took the "Physical Review" and marked an article for his student to analyze. The articles were from absolutely different areas of physics, and it was necessary to prepare in a week (see: "Landau was the only encyclopedist for the entire time of science development", 6).

The Soviet science structure and organization has been deployed in sufficient detail on the modern network platforms. Basically, this is about the Academy of Sciences, nuclear and space projects, mainly in the aspect of relationship between the scientific environment and the government. A number of materials are related to the anniversary of the creator of the hydrogen bomb and human rights activist *Andrei Sakharov*.

Semantic aspect of the media representation.

Soviet science was called to replace the repressed religion in the public consciousness. During the Soviet era, scientist itself was a cult person, and the space (cosmos) was

something like new religion. Science popularization was interpreted as a scientist's social duty (at the same time, certain standards of popularization were formed), and the status of scientific knowledge was cultivated as elite and closeness (closed to the uninitiated) at the same time.

The Soviet media propaganda entire system (newspapers, radio, TV) constantly broadcasted thesis about priority position of science into the Soviet society. The role of science in the modern Russia, unfortunately, cannot be even compared with the Soviet era, although efforts have been made to fix the situation. The first human flight into the space still interpreted as one of the main factors in uniting the nation. Anniversary of the first human to fly to space is widely represented, and *Yuri Gagarin* himself becomes a media celebrity.

The popularization of knowledge was carried out through the press: there was a science heading in almost in each edition (newspaper, magazine). In the 1960s – the first half of the 1980s, the iconic brand of Soviet popular science journalism, the magazine with millions of circulations, "Science and Life", has been extremely popular among the Soviet intelligentsia.

The surviving Soviet popular science magazines have lost their influence, and their history mostly hadn't represented by the digital popular science (one of the notable exception: "Novelties of inventions in the magazine 'Science and Life', 1893", 2), but dissertations about them have being written from time to time.

The symbolic status of science was ensured by cinema (movies) and fiction. Initially, in the early 1930s, the Soviet communist party leaderships formed an order for the science coverage by fiction (literature). The scientific agenda became organic very quickly to the Soviet literature and art. Avant-garde art corresponded to the avant-garde science, searching immortality for the political elite, rejuvenation experiments, the abolition of the sexes, and so on.

Science as metatext associated also with the popular culture, the most in demand within modern digital popular science.

Pragmatic aspect of the media representation.

Science was almost the only single legal channel for avoiding state ideology. It is really noteworthy that *Khrushchev's* "thaw", began with the popular science literature. Sci-Fi provided social engineering. There was a real cult of *Ivan Efremov's* novel, named "Andromeda Nebula", the exploration of the galaxy has been closely connected with the planetary communist society idea. At this time, against the background of de-Stalinization, a return was made to the ideas of the theorists of the socialist state (*Lenin, Trotsky*), who called for the world communism. The construction of world communism was transferred to the open space, or the ocean depth. The unity with movies and literature (fiction) is especially important in the science mediatization process. Famous *Andrei Tarkovsky* movie "Stalker" was based on brothers *Strugatsky* novel, as "Solaris" based on the novel of the same name by the Polish science fiction writer and futurist *Stanisław Lem*. Brothers *Strugatsky* are some of the most frequently mentioned characters on the digital platforms (see: "Revived wolds of *Strugatsky* brothers", 3).

Scientific publicism role was extremely high during *Gorbachev's* perestroika (reconstruction). Leading social and humanitarian scientists, academicians spoke publicly about ideological and economic renewal of the country.

The ideas of social responsibility of the science and scientists in the situation of networked civilization and artificial intelligence establishment are becoming the most popular.

A number of materials are devoted to the futurological projects, including social engineering, which have arisen as a consequence of cybernetic development (see: "How dissidents used cybernetics tried to change the world", 1).

Results

Science enjoyed absolute authority into Soviet society. The level of the Soviet science mediatization can be described as really deep, and success efforts have been contributed by:

- 1) Comprehensive state support for the scientific knowledge promotion, which determined the comprehensive communication channels;
- 2) Focusing not on scientific PR, or monetization, but on the audience's scientific worldview development;
- 3) Symbolic potential of science was based on its unity with literature and art;
- 4) Correlation with the public agenda and moral searches of the Soviet intelligentsia.

Soviet experience in the science mediatization demonstrates a constructive meaning that supplemented (and is some ways – not duplicate) the Western system. In today's pandemic context and closed societies, this experience becomes relevant again. Pandemic has revealed significant problems in communication between science and society. And this problematic is connected, among other things, with the superficial, not deep nature of science mediatization general process, what is clear on the example of the Soviet science digital grey. Political and ideological agenda replaces meaningful conversation about science as such.

References

1. *Arzamas*. Accessed January 2022. <https://arzamas.academy/materials/2277>
2. *Diletant.media*. Accessed January 2022. <https://diletant.media/articals/37418872>
3. *Diletant.media*. Accessed January 2022. <https://diletant.media/articals/32145157>
4. *Journalism History and Digital Archives*. 2021. Ed. by Henrik Bødker. Routledge (Taylor & Francis).
5. *Lazarevich, E.A. 1981. Populyarizatsiya nauki v Rossii [Science popularization in Russia]*. Moscow: Moscow University Publishing.
6. *PostScience (PostNauka)*. Accessed January 2022. <https://postnauka.ru/talks/87126>

Zines as Nonbinary Objects and Questions of Privilege

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

Introduction

Zines exist within the hybrid recesses of cultural consciousness. They are subcultural relics of the eras prior to the internet and they are knowledge infrastructures that overtly disrupt mainstream cultural perspectives. They also function as boundary artifacts that bridge gaps between analog print protocols and digital manifestations of data such as PDFs and hyperlinks. Nicolini et al. (2012) observe that “we look at objects as boundary devices but also as epistemic things, objects of activities, and infrastructures” (p 5). Zines as critical data artifacts communicate subcultural values and rebellious ideologies. Zines are also hybrid creations that traverse the boundaries between digital and analog as well as the social and scholarly.

Zines as “epistemic things” (Nicolini et al. 2012 p. 5) emerge as cultural representations but also adapt to serve as attitudes about “the world that leads to the deposition of knowledge in a form from which others can draw” as Rheinberger (2005 p. 409) commented. Zines traditionally are laypersons’ objects, with content that is created for and consumed by other laypersons. They possess rebellious potential in that not only do they represent the ideas which they espouse and the cultures which generate these notions, but they communicate ideologies across social margins. They are traditionally laypersons’ objects, with content that is created for and consumed by other laypersons.

Zines are social objects too, and they draw their influences from underground and nonlinear processes that make them “distinct from other modes of relating to the world” (Rheinberger 2005 p. 409). The “data journeys” (Longino 2020 p. 391) of zines are mutable messy and multiple; as social objects they do not pursue linear trajectories and as knowledge infrastructures they are porous as they evolve through the participatory responses, reactions and reflections of the individuals who read them. R. Clark Parsons (2017) has indicated that “Over the last decade, zines, or self-published booklets ranging dramatically in style and content, have made a resurgence. It is impossible to estimate the number of contemporary zinesters in the United States, whose subversive, hodgepodge texts are not catalogued in the Library of Congress or issued ISSNs... The humble do it yourself (DIY) zine perseveres in spite of, but perhaps more accurately, *because* of the meteoric rise of blogging and social media platforms” (p.2). Zines occupy the digital and the analog terrains simultaneously. However, R. Clark Parsons (2017) also interjects that “the polarization of zines versus blogs also precludes a more nuanced framework that positions zine-making as a feminist practice working in conjunction with digital media, distinct but symbiotic discursive strategies for coping with structures of power that privilege some bodies while marginalizing others within the public sphere” (p 4). Zines are social objects too, and they draw their influence from underground and nonlinear processes that make them disruptive and disorderly.

Furthermore, data themselves are messy and mutable; zines capture this reality through their lack of formalized classification, indexing and bibliographic archives during their creation and distribution. Longino (2020) expands on this, concluding that it is “a naïve fantasy that data have an immediate relation to the phenomena of the world, that they are ‘objective’ in some strong, ontological, sense of that term, that they are the facts of the world speaking directly to us” (p. 391). Zines as boundary objects are subjective, social

and subversive data artifacts that not only demonstrate subcultural ideologies, but collaboratively create communities of practice through shared interest and perspectives. Clark- Parsons (2017) supports this through her observation that “Zine-making as an accessible DIY media practice that operates outside of both marketplace logic and sociopolitical constraints, enables the invention and circulation of counterdiscourses that might otherwise find no outlet within the commercial media landscape” (p 9).

Zines such as *Major Threat* (a punk rock boundary object designed by academics and educators to reach laypersons) is a data artifact that loudly proclaims how zines occupy digital and physical communities and resist conventional classification schemes. Kemp, founder of *Major Threat*, tells the readers that “if you find a copy of this online, please share it. Print a couple. Email to friends. Post it anywhere and everywhere. Help this movement to improve education grow” (Kemp 2020 p. 2). This particular zine uses the viral philosophical practices of the digital realm to encourage widespread meme awareness in print formats and in online platforms as well. *Major Threat* promotes the mores of punk literati to dismantle traditional academia and simultaneously calls for rebellion from laypersons to effect this change from a place of privilege.

The embodied contradiction in this zine-as-data-thing deftly reveals how zines resist binary classifications as simply literature, ephemeral objects or digital representations of countercultural insurrection. Grey literature supports rebellious data things as it calls attention to contradictions as reflexive spaces where people can delve into transgressive knowledge infrastructures built through collaborative, social and subjective interactions within varied communities. Zines are one crucial example of the sundry possibilities to engage in conversations beyond the cross disciplinary and ask difficult questions about accessibility, privilege and diversity.

Zines as cultural artifacts and librarians

Creasap (2014) has argued that “Zines occupy a middle ground between traditional research papers or essays and Web-based media such as blogs. Unlike research papers, zine style is decidedly informal. Images are hand drawn or cut and pasted by hand..The informal, creative and participatory character of zines...unlike blogs, zines are physical objects that can be held and passed from person to person by hand” (p.155).

Yet as demonstrated by the aforementioned zine *Major Threat*, zines are also digital objects that circulate in online environments. Zines as cultural artifacts are constructed as countercultural data objects; as Creasap (2014) also finds “Zines share commonalities with independent media of earlier women’s movements, such as scrapbooks, pamphlets and manifestos” (p. 157). The scholar C. L. Weida (2013) argues that this countercultural significance of zines has antecedents in other types of DIY objects when she observes that “Zines may be seen as an extension of genres like artist sketchbooks, chapbooks, surrealist games, and manifestos of art history” (p. 68). C.C. Bagelman and J. Bagelman (2016) comment on this further by linking zines to “philosophical movements like Surrealism used small runs of self-published material, decorated in collage and bricolage, as a forum for ideas” (p. 366). Zines have also, according to N. Nijsten (2016) “evolved to include punk fanzines from the 1970s onward, and since the 1990s more and more women got involved” (p. 414). Zines as symbols of countercultural and grassroots community values within their historical antecedents continue to fascinate information professionals and educators. Many librarians are fascinated by zines and S. Thomas (2018) has commented on how librarians in particular “have expressed interest in collecting and teaching zines

and regularly seek opportunities to co-teach or present to classes. Librarians can also assist with zine assembly and reproduction outside the classroom. A librarian may know of other faculty teaching with zines, of campus resources and contacts for planning a zine event and of relevant special collections in the library” (p. 750). Librarians do have access to knowledge infrastructures that include special collections and the opportunities to engage with communities and classrooms in various ways that students and faculty might not be able to do.

Du Laney, Maher and Schindler (2020) have argued that “The zine format becomes a vehicle that integrates key skills such as research techniques, critical information literacy and concise argument synthesis with course-specific learning outcomes” (p. 12). Lymn (2013) expands this to include librarians as potential countercultural producers of knowledge infrastructures as well as information allies when she constructed “librarians as insider ethnographers” (p. 1) within their communities. She directly confronts conventional perspectives of librarians as custodians of traditional mores when she states: “There is the sense that the librarians don’t participate in DIY and grassroots communities; but they do” (Lymn 2013 p. 4). S. Britton (2018) has expressed the odd status of zines in library spaces as well as the difficulties presented in terms of access, authority, and agency when zine values and countercultural resistance encounter institutional barriers and metadata management. She comments that “Some zine makers are reticent for their zines to be part of library collections, particularly institutional libraries as the controlled, owned, nature of something like a library collection seems to be the antithesis of many of the fundamental ideas that are central to zine making: independence, estrangement from mainstream culture, and the zine as an ephemeral object” (Britton 2018 p. 5).

Zines as NonBinary Objects

Zines as artistic expressions, sentimental confessions, cultural critiques and community manifestos establish themselves as nonbinary objects. They disrupt digital spaces through an insistence on print media as a way to spread memes of resistance and revolt. They disrupt analog spaces through an insistence that community and creativity can expand and grow through digital connections. Alison Piepmeier observed in 2008 that zines are “what Gregory Sholette terms ‘Dark Matter’, work that functions outside of and is therefore invisible to the established art world and to academic scholarship” (p. 218). C.L. Weida (2013) indicates that the zine as a nonbinary objects demystifies knowledge infrastructures as hidden from the layperson; rather it builds infrastructures through creativity and community: “..making a zine does not require training, initiation, or education as a prerequisite-a zinester is simply a person who creates a zine.” (pp. 76-77). Zines as hybrid nonbinary social objects “explicitly explore the intersection of narrative and materiality” (Poletti 2008 p. 87).

M.R. Bold (2017) also discusses why zine making has grassroots outreach through digital spaces and why it is critical for zine communities to be diverse, active and visible. “Zines are significant because they offer the opportunity for connection, community and networking between those interested in these diverse topics. Despite this..diversity in race, class, and age are underrepresented in the zine community.” (p. 215). However, “Zines and DIY/self-publishing have helped to change the way that consumers engage with content: turning passive consumers into active cultural producers” (Bold 2017 p. 219). Zines as nonbinary data artifacts create agency and have the potential to generate communal ideologies that challenge conventional boundaries and barriers to laypersons because as A. McNutt (2021) observed “Zines allow for personal expression in a manner

that most other mediums do not. This is likely due to history with individualistic or alternative cultures...as well as the amorphous nature of their form-typically any approach to their design or content is valid. The freedom of personal expression is thus not bound by the restrictive more found in other forms” (p. 3). Zines are one crucial example of the sundry possibilities to engage in conversations beyond the cross disciplinary and to ask questions about agency, privilege, accessibility and diversity within cultures that are valued and created by laypersons.

Works Cited

- Bagelman, J. J., & Bagelman, C. (2016). Zines: Crafting change and repurposing the neoliberal university. *ACME: An International Journal for Critical Geographies*, 15(2), 365-392.
- Bold, M. R. (2017). Why diverse zines matter: A case study of the people of color zines project. *Publishing Research Quarterly*, 33(3), 215-228.
- Britton, S. (2018). What we do, is (still) secret? Collection, care and accessibility of zines in UK collections. *Art Libraries Journal*, 43(2), 72-76.
- Clark-Parsons, R. (2017). Feminist ephemera in a digital world: Theorizing zines as networked feminist practice. *Communication, Culture & Critique*, 10(4), 557-573.
- Conn, V. *Major Threat: Punk Rock Academia 9 (A 'Zine)*. 2020. Drew Kemp issue #1.
- Creasap, K. (2014). Zine-making as feminist pedagogy. *Feminist Teacher*, 24(3), 155-168.
- Du Laney, C., Maher, M., & Schindler, A. (2020). The Librarian's Guide to Zines for Classroom and Community.
- Fox, V., McElroy, K., Vachon, J., & Wooten, K. (2018). Each according to their ability: Zine librarians talking about their community.
- Gray, E. M., Pollitt, J., & Blaise, M. (2021). Between activism and academia: zine-making as a feminist response to COVID-19. *Gender and Education*, 1-19.
- Hays, A. (2020). A Citation Analysis about Scholarship on Zines. *Journal of Librarianship and Scholarly Communication*, 8(1).
- Lymn, J. (2013). The librarian-as-insider-ethnographer. *Journal of Library Innovation*. 4(1), 1-10.
- Longino, H. E. (2020). Afterword: Data in Transit. In *Data Journeys in the Sciences* (pp. 391-399). Springer, Cham.
- McNutt, A. (2021, October). On the Potential of Zines as a Medium for Visualization. In *2021 IEEE Visualization Conference (VIS)* (pp. 176-180). IEEE.
- Nicolini, D., Mengis, J., & Swan, J. (2012). Understanding the role of objects in cross-disciplinary collaboration. *Organization science*, 23(3), 612-629.
- Nijsten, N. (2016). Write it yourself?: Feminist perzines as participatory playgrounds. *Tijdschrift voor Genderstudies*, 19(4), 413-431.
- Piepmeyer, A. (2008). Why zines matter: Materiality and the creation of embodied community. *American Periodicals*, 18(2), 213-238.
- Poletti, A. (2008). Auto/assemblage: Reading the zine. *Biography*, 85-102.
- Rheinberger, H. J. (2005). A reply to David Bloor: "Toward a sociology of epistemic things". *Perspectives on Science*, 13(3), 406-410.
- Thomas, S. (2018). Zines for teaching: A survey of pedagogy and implications for academic librarians. *portal: Libraries and the Academy*, 18(4), 737-758.
- Weida, C. L. (2013). Feminist zines:(Pre) Occupations of gender, politics, and DIY in a digital age. *Journal of Social Theory in Art Education*, 33(1), 67-85.

The endless life of OA Journals from myth to reality: the survey on present status of vanished OA journals in Iran and future prospect

Hamideh Memari Hanji, ISSN National Center, Iran

Abstract:

Open access journals in Iran were created in line with the open access movement in the world and their number is increasing day by day. In Iran ROAD is recognized as an important resource for identifying and tracking open access journals for publishers, users and researchers. In monitoring the Iranian open access journals of ROAD database, we recognized that the URL of some of these journals is not active. As we know there is no guarantee for long-term access to open access journals on the web & Iran doesn't have Internet archives, the importance of investigating this issue becomes more. The primary aim of this study for the first time is to explore the phenomenon of vanishing OA journals in Iran on ROAD by tracking all Iranian open access journals in this database in the period from 2010 - 2020 in all subject area such as social science, applied science, medicine & etc. the journals were sorted by year, and in each year they divided into three categories base on type of publishers: academic publishers, non-academic publishers and self-publishers. After checking the ROAD URLs, if the address is not active, title and URL of the journals were searched in all kinds of search engines and Wayback Machine. Our result shows that the highest percentage of vanished journals belong to the category of self-publishers and non-academic publishers, respectively, and academic journals have the lowest percentage of vanished. Surveys show that between 2010 and 2020, a total of 110 open access journals were vanished, including 38 self-publishers, 25 academic publishers and 47 non-academic publishers.

Keywords: Iranian open access journals, Vanished OA, ROAD, Iran ISSN National Centre, ISSN international Center

Introduction

Scientific journals are one of the main channels of communication between researchers and experts. Researchers publish the results of their research in relevant and reliable scientific journals. In many countries, the credibility and effectiveness of researchers is evaluated based on the quantity and quality of their articles published in scientific journals. The history of the invention of the first scientific journals dates back to the 17th century¹

In the past, people had access to scientific journals in print and through the payment of a subscription fee, but in recent years, many of the features of scientific journals have changed. The publication of online magazines, which began in the early 1990s with the development of the World Wide Web, is considered to be the most influential of these developments. Also, the expansion and dissemination of free and unrestricted research outputs in the form of the Open Access Movement. Today, some of the scientific articles in scientific journals are published in the field of science beyond open access. It's been just over a decade since the concept of Open Access (OA) first captured the attention of the scientific and scholarly research community, bringing with it the promise and potential of a shining new digital landscape, in which knowledge is freely shared and freely used, and the pace of scientific discovery is accelerated for the benefit of all. ²

The Budapest Open Access Initiative (BOAI) is a public statement of principles relating to open access to the research literature, which was released to the public on February 14, 2002. It arose from a conference convened in Budapest by the Open Society Institute on December 1–2, 2001 to promote open access which at that time was also

known as Free Online Scholarship. This small gathering of individuals is recognized as one of the major defining events of the open access movement. The text of the initiative was translated to 13 languages.

On the occasion of the 10th anniversary of the initiative in 2012, the ends and means of the original initiative were reaffirmed and supplemented with a set of concrete recommendations for achieving open access in the next 10 years.³

In 2012, the ISSN International Centre, under the guidance of its Governing Board and largely inspired by the ventures mentioned above, posited that it could play a useful part in the promotion of Open Access scholarly resources. With the support of the 89 national centers which comprise its network, the ISSN International Centre is indeed in a good position to provide an overview of the development of OA scholarly resources worldwide. With backing from UNESCO’s Communication and Information Sector, the ISSN International Centre opened in 2013 a web service called the Directory of Open Access Scholarly which is fed by national ISSN centers supplying their bibliographic records that are further processed to be made available on ROAD in various formats including RDF. A unique feature of ROAD lies in the provision of global statistics thus allowing users to monitor the development of OA resources across the globe.⁴

e-journals in Iran

Along with the emergence of electronic journals in the world, Iranian e- journals also grew day by day. Statistics show that the number of e-journals is increasing every year compared to print journals.

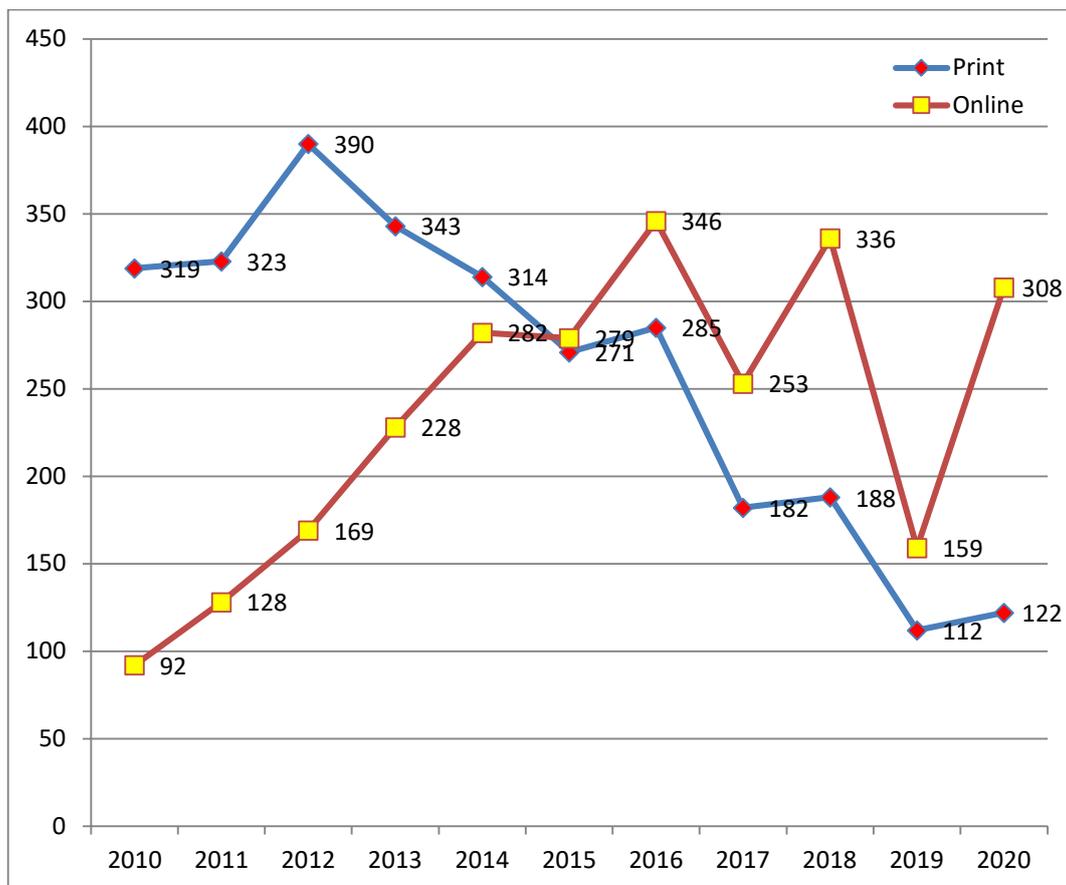


Figure 1- Increase in the number of e-journals 2010 - 2020⁵

Moreover, special features of open access publications have led many publishers to move towards the open access of their information resources. This process is accelerating and the number of open access journals is increasing every day. The short distance between the completion of the research and its publication and the possibility of wider access of

the scientific community to open access journals has caused more success and attention to it.

Open Access background in Iran

Since 2013, Iran ISSN National Center along with the establishment of ROAD, like other ISSN network centers, start cooperated with ROAD. According to ROAD activity report for 2018, Iran is one of the top ten contributing countries in OA publications.⁶

Table 1. Top 10 participating countries (ISSN National Centres

Indonesia	4920
France	2912
India	2410
Turkey	1749
United Kingdom	1666
United States	1655
Iran	1456
Brazil	1286
Poland	1219
Spain	926

ROAD statistics until the end of 2020 shows:

- 2163 journals are published as open access.
- 84 journals are archived in Keepers Registry
- 66 titles have been Ceased

The subject coverage of open access journals is as follows:

- SOCIAL SCIENCES (1408 Title)
- APPLIED SCIENCES. MEDICINE. TECHNOLOGY (934 Title)
- MATHEMATICS. NATURAL SCIENCES (212 Title)
- RELIGION. THEOLOGY (148 Title)
- LANGUAGE. LINGUISTICS. LITERATURE (127 Title)
- THE ARTS. RECREATION. ENTERTAINMENT. SPORT (90 Title)
- PHILOSOPHY. PSYCHOLOGY (87 Title)
- GEOGRAPHY. BIOGRAPHY. HISTORY (87 Title)
- SCIENCE AND KNOWLEDGE. ORGANIZATION. COMPUTER SCIENCE. INFORMATION. DOCUMENTATION. LIBRARIANSHIP. INSTITUTIONS. PUBLICATIONS (68 Title)

The number of journals that are indexed in indexing databases by type of database is as follows:

- ROAD (2163)
- DOAJ (451)
- CROSSREF (380)
- SCOPUS (129)
- CABABSTRACTS (84)
- THE KEEPERS (84)
- GLOBALHEALTH (75)
- PUBMED (54)
- PUBMEDCENTRAL (32)
- ECONLIT (9)
- MIRABEL (7)
- MEDLINE (4)
- PSYCHINFO (3)
- CIRAD (OU PUBLIER) (3)
- GEOREF (3)
- LINGUISTICS (1)

Methodology

- At the first we define a “vanished” OA journal as a journal that published at least one volume as immediate OA after which production ceased, and the journal, together with the published full-text documents, disappeared from the web.⁷ In cases, individual issues of the journals are still exist on the web, through local and international indexing open access database, we do not count the journal as a vanished journal. Also, if a URL is searched in the wayback machine and full-text volumes and articles are found, it will still not be considered as vanished journals.
- The research population includes all Iranian Open access journals in the ROAD, for identifying vanished OA journals in this research we focus on vanished open access journals in ROAD database & we selected all number of Iranian OA journals between the years 2010 - 2020 in all subject area such as social science, applied science, medicine & etc.
- Journals were sorted by year, and in each year we divided them into three categories by type of OA publishers: academic publishers, non-academic publishers and self-publishers.
- We focus on journals instead of articles for methodological reasons then we start to examined ROAD URLs, if the journals URL is not active, we searched both title and URL of the journals in search engines and Wayback Machine. through various searches we found that in some cases only the URL of the magazine has changed and the magazine is still active and current, so we were recorded the number of these records by category as a magazine with changed URL and also listed journals that were completely vanished by category. After a month URL of vanished OA Journals In the first review have been re-checked to make sure whether they were really vanished or not.

Findings

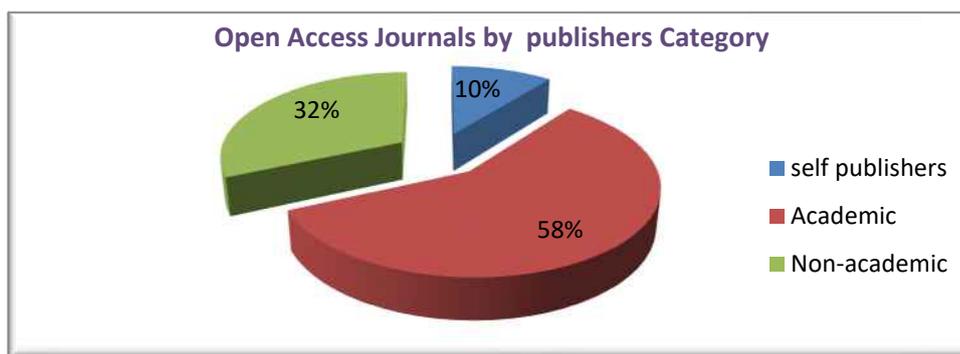


Figure 2 - Open Access Journals by publishers

Analysis of open access journals published between the years 2010-2020 showed that these journals were published by publisher type as follows:

- 10.63% self-publishers
- 57.65% Academic journals
- 31.71% non-Academic

Table 2 - Review statistics of open access journals by year

Year	Sum	Publishers											
		Self-Publisher				Non-person publishers							
						Academic publishers				Non-Academic publisher			
		Vanished	Changed Address	Current	Total	Vanished	Changed Address	Current	Total	Vanished	Changed Address	Current	Total
2010	40	0	0	1	1	0	2	28	30	0	0	9	9
2011	57	1	0	5	6	0	3	39	42	0	0	9	9
2012	61	0	0	0	0	0	5	34	39	3	3	16	22
2013	120	3	0	7	10	1	3	69	74	5	3	31	36
2014	229	3	1	17	21	6	10	112	128	6	1	83	90
2015	250	6	2	10	18	5	9	128	142	6	4	75	90
2016	312	9	5	25	39	3	6	177	186	6	3	78	87
2017	317	6	5	15	26	4	12	188	204	11	8	68	87
2018	324	9	0	143	52	2	6	163	171	5	2	105	101
2019	159	1	0	18	19	2	3	71	76	3	0	61	64
2020	284	0	0	38	38	4	0	151	155	2	0	89	91

- ✓ A review of 2010 journals found that 75% of them were academic, 22.5% were non-academic, and 2.5% has self-publishers. 6.6% of academic publishers have changed URL.
- ✓ Among open access journals in 2011, 68.73% were academic, 78.15% were non-academic and 52.10% were self-publishers. 7.14% academic publishers have changed their address, 16.66% of self-publishers have been vanished and could not be retrieved in internet search.
- ✓ In 2012, 63.99% of open access journals have academic publishers and 36.65% of them had non-academic publishers and there were not self-publishers. The study showed that 12.82% of academic journals and 13.632% of non-academic journals have changed

their addresses, and finally 9.09% of journals with non-academic publishers have been vanished.

- ✓ A review of 2013 shows that 61.66% are academic publishers, 30% are non-academic publishers and 8.3% are self-publishers.

In this year 2.7% academic publishers, 11.11% non-academic publishers and 30% self-publishers have vanished & 4.05% academic publishers, 8.33% non-academic publishers Have changed the URLs.

- ✓ There were 55.89% journals with academic publishers, 39.30% of with non-academic publishers and 4,080% of self-publishers in the year 2014. Of these, 4.68% academic publishers, 1.11 non-academic publishers, and 4.76% self-publishers journal have vanished. 7.8% academic publishers, 7.77% non-academics, and 14.28% self – publishers have changed Address.
- ✓ In 2015, 250 open access journals were published, of which 56.8% were academic, 34% non-academic and 9.2% were self-publishers.

3.52% of academic journals, 6.66% non-academic journals & 33.33% self-publishers have vanished. In 6.33% academic journals, 4.44% non-academic journals & 11.11% URL change observed.

- ✓ Among the 2016 open access journals, 59.61% academic, 27.88% non-academic and 12.5% self-publisher have observed.

1.61% academic journals, 6.87% non-academic journals & 12.82% self-publishers have vanished. URL change was observed in 2.67% academic 3.44% non-academic and 10.25% of self – publishers.

- ✓ In 2017, the number of open access journals reached 317, 64.35% were academic, 27.44% were non-academic and 7.95% were self-publishers. 1.47% of academic journals, 13.79% of non-academic journals and 23,076% of publishers have vanished, respectively.
- ✓ Among the open access journals in the year 2018, 52.77% were academic, 31.7% were non-academic and 16.04% were self-publishers. After reviewing the URLs of magazines published in this year, it became clear that 4.95% non-academic publishers & 17.3% of self-publishers have vanished in this year. 3.50% journals with academic publishers, 1.98% non-academic publishers & 3.84% of self-publishers had changed URL.
- ✓ Among published open access journals in the year 2019, 47.69% were published by academic publishers, 40.25% by non-academic publishers and 11.94% by self-publishers. 2.64% open access journals by academic publishers, 4.68% by non-academic publishers and 5.26% by private publishers have vanished.
- ✓ Among 284 open access journals in the year 2020, 54.57% belongs to academic publishers, 32.04% non-academic & 13.38% were self-publishers. In this year just 1.29% of journals with non-academic journals vanished.

Results

Studies show that the highest percentage of vanished journals belong to the category of self-publishers and non-academic publishers, and academic journals have the lowest percentage of vanished.

Number of vanished OA by publishers category

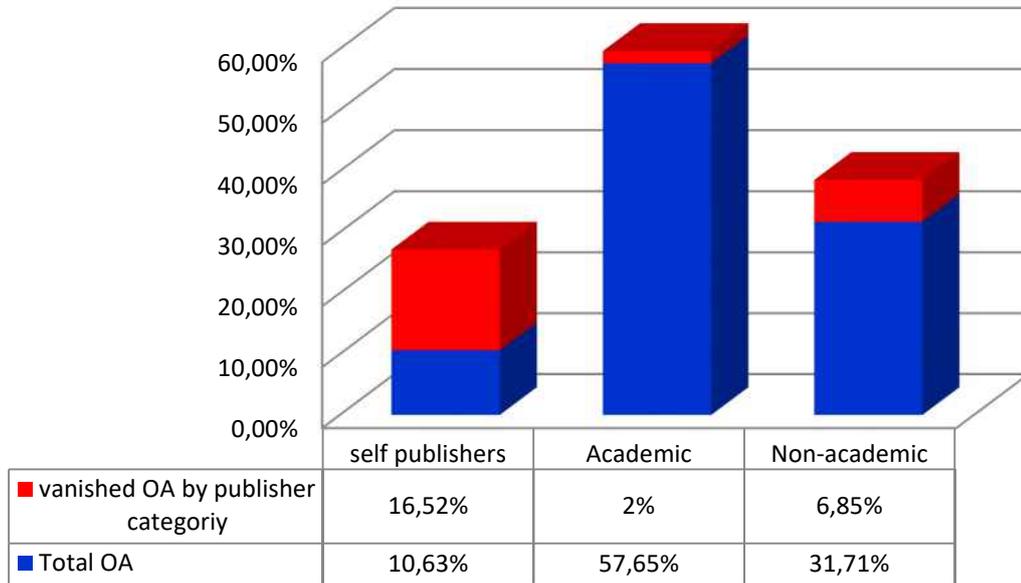


Figure 3 - Number of vanished OA by publishers category

Academic journals in 2014, non-academic journals in 2017 and self-publishers in 2015 had the most vanished OA journals. Surveys show that between 2010 and 2020, a total of 110 open access journals were vanished, including 38 self-publishers, 25 academic publishers and 47 non-academic publishers

vanished OA by publishers category

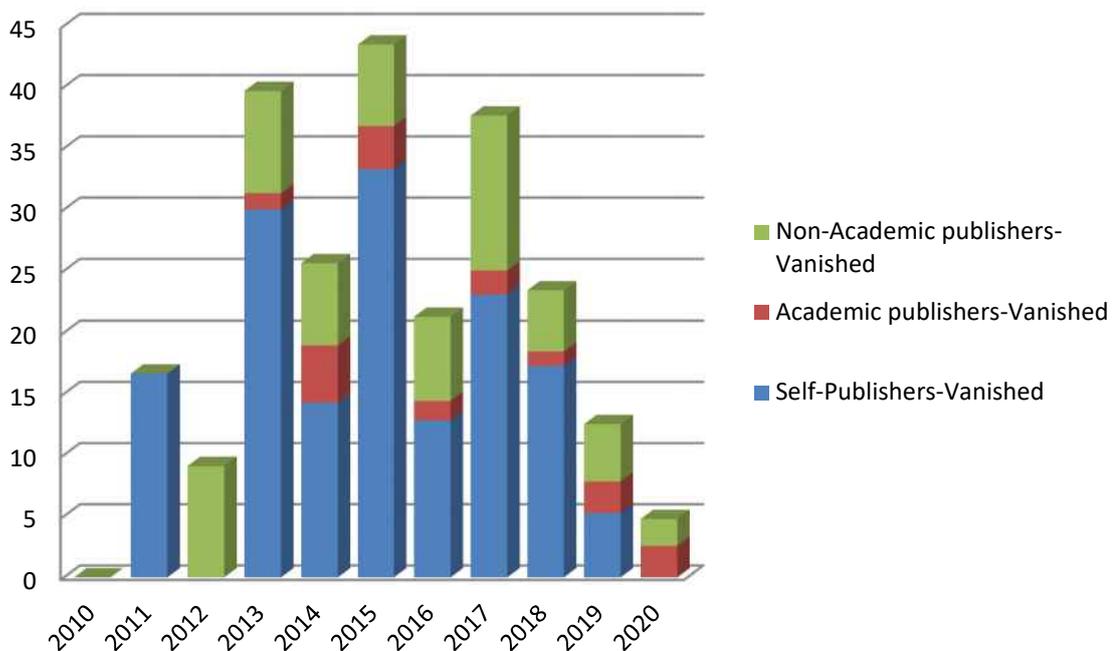


Figure 4 - vanished OA by publishers category

The most URL changed belongs to Self-publishers and non-academic publishers.

URL Changed by Publishers Category

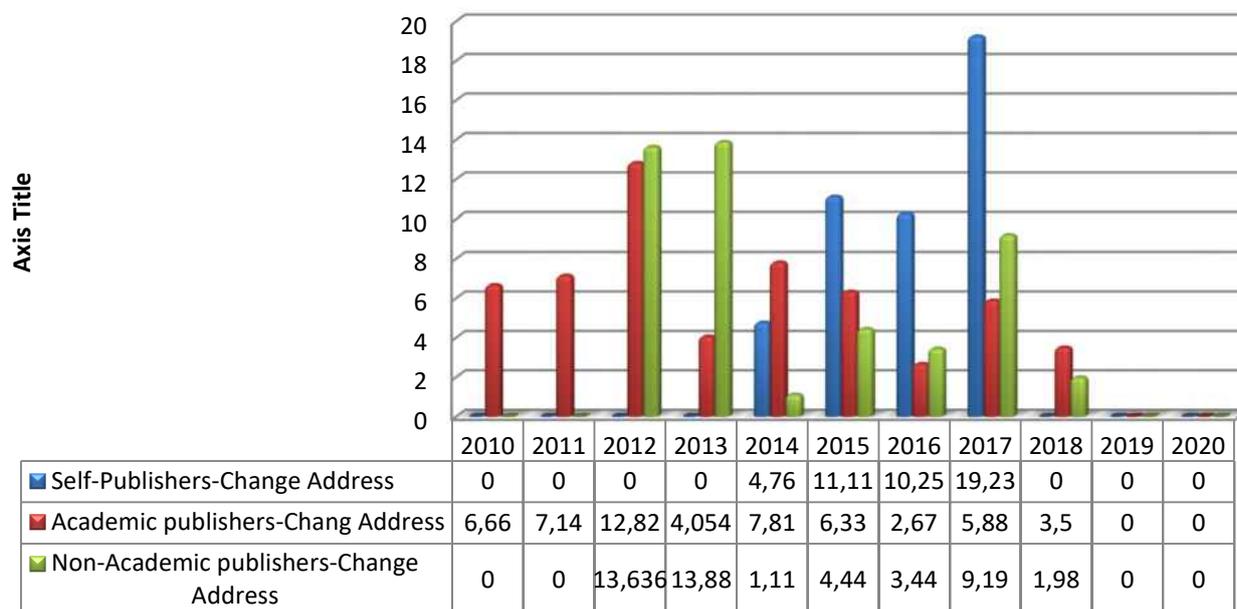


Figure 5 - URL Changed by Publishers Category

Conclusion

- ✓ Preserving OA journals is big financial challenge for publisher so it seems the highest percentage of vanished OA journals belong to the category of self-publishers.
- ✓ Academic journals not only have less financial challenges but also have a more consistent policy.
- ✓ The criteria for assigning ROAD code to OA journals are updated annually and announced by the ISSN International Center to the ISSN networks, the Iran ISSN Center also follows these criteria, It seems that the use of these criteria has helped to reduce the number of vanished OA in recent years.
- ✓ None of the vanished journals were indexed in any of the international or local indexing databases
- ✓ The vanished journals were completely removed from the web and it was not possible to contact them to find out how many issues had been published.
- ✓ keepers Registry can protect e-journals which are “at risk of loss” and need to be archived.⁸
- ✓ In fine, implementing the National Web Archives in Iran is essential.

¹ Bill Cope, Angus Phillips. The Future of the Academic Journal. Elsevier Inc.

² Heather, Joseph, The Open Access Movement Grows Up: Taking Stock of a Revolution

³ https://en.wikipedia.org/wiki/Budapest_Open_Access_Initiative

⁴ Béquet, Gaëlle. Questionable Practices in Scholarly Publishing: the Stance of the ISSN Network .IFLA Library , 2016

⁵ Portal.issn.org

⁶ <https://www.issn.org/wp-content/uploads/2019/09/ROAD-statistics-for-2018.pdf>

⁷ Laakso, Mikael, Matthias, Lisa. Open is not forever: A study of vanished open access journals. Journal of the Association for Information Science and Technology Volume 72, Issue 9 p. 1099-1112

⁸ <https://keepers.issn.org/keepers-registry>

Burning Grey: The Worldwide Influence of a Locally Published Grey Literature

Vince Ervin V. Palcullo, Donna May C. Rivera, Via Marie F. Dumenden,
Ma. Cynthia Peleña, Central Philippine University, Philippines
Joy Geromiano and Daryl Superio, Aquaculture Department,
Southeast Asian Fisheries Development Center, Philippines

Abstract:

Rice is a major staple crop in the Philippines, which produces a large amount of rice each year. On the other hand, rice production generates biomass waste in the form of rice husks. In consideration of the rice husks' potential as a biomass energy source and support the development of the rice husk gas stove technology in the Philippines, Engineer Alexis T. Belonio had published a grey literature (GL) entitled "Rice Husk Gas Stove Handbook." Hence, to fully understand the impact of the technology, a citation analysis was conducted. The works that cited the GL were retrieved from Google Scholar using Harzing's Publish or Perish software. The complete title of the GL was used as keywords for the search string. Results revealed that the GL was cited by 155 literatures written by 398 authors from 31 countries. Seventy-five (48.4%) of the 155 literatures were cited 797 times. The results showed the wide utilization and the impact of locally-published GL, thereby confirming GL's value in research and development.

Introduction

Rice production is one of the most essential economic activities in the world (GRiSP, 2013). It is the staple food of more than 50% of the world's population, with more than 700 million tons produced annually (CGIAR, 2019). Rice is primarily a foodstuff and continues to be a major food staple in Asia, Latin America, the Caribbean, and increasingly in Africa, and the global rice production is expected to grow by 58 mt to reach 567 mt by 2030 (OECD/FAO, 2021). In the Philippines, rice is the staple food and is one of the most important staple crops in the country. As a matter of fact, Philippines ranks as the world's eighth-largest producer of rice (GRiSP, 2013). In 2020 alone, a total of 19.44 million metric tons (MT) of rice was produced (Agoot, 2020), which also made a significant amount of agricultural waste, such as rice husks. Rice husks or rice hulls are the hard protective covering of rice grains and are the by-product of rice production during the milling process. An estimated 2 million tons of rice husk are produced annually in the Philippines (Fung and Jenkins, 2003; Belonio, 2005). Rice husk is considered as the most abundant agricultural by-product in the country (Vinluan, 2002), and it holds a great potential to be a biomass energy source that could help address the population's increasing fuel needs according to Philippine Rice Research Institute as stated in the paper of Simeon (2016).

According to Lim et al. (2012), huge dependence on conventional and fossil fuels pose a significant problem in fossil fuel depletion, climate change, and environmental protection. Additionally, though fossil fuel is predicted to remain the dominant energy source by 2030 (Lim et al., 2012), Mofijur et al. (2019) argued that it would be incapable of supporting the global energy demand in the future due to its limited reserves. Many developing countries generally use Liquefied Petroleum Gas (LPG) as conventional fuel source for cooking (Punin, 2020). In the Philippines, conventional source of fuel such as the LPG is commonly used by many households (Belonio, 2005). However, the volatile oil price and growing emphasis on environmental conservation has led to the development, seeking, and utilization of alternative energy sources (Lim et al., 2005). Thus, the Department of Agriculture in the Philippines and the International Rice Research Institute developed the rice husk gas stove technology in 1986. Aside from its financial viability, utilization of rice husks as a biomass energy source through this technology also offers environmental sustainability by reducing the rice husk disposal problem (Belonio, 2005; Kumar, 2013).

Also, it has less carbon dioxide emission than traditional fuel sources in cooking, therefore “contributing to carbon sequestration for greenhouse gas mitigation” (Simeon, 2016).

Hence, to contribute to the development of the technology, Central Philippine University in Iloilo City, Philippines, through the leadership of its faculty member Engineer Alexis T. Belonio, developed a rice husk gas stove in 2003 (Strauss, 2016). Furthermore, to fully disseminate the technology, the University published in 2005 the "Rice Husk Gas Stove Handbook," a grey literature (GL) that is being used as one of the prime references in sustainable energy sources research.

Grey literatures are information produced in print and electronic formats not controlled by commercial publishing, where publishing is not the primary activity of the producing body (Tillett & Newhold, 2006). Grey literature is usually produced by associations, academic institutions, research institutions, libraries, societies, etc. common types are theses, unpublished documents, conference proceedings, datasets, reports, working papers, etc. (Mason, 2012; Schopfel & Farace, 2009).

GL’s value in bringing unnoticed, unexplored, and difficult to acquire information from different facets of society contributes to the development of a nation through its literary output (Gul, Shah, Ahmad, Gulzar, & Shabir, 2021). However, GL’s usability and utilization are faced with issues of discoverability due to the locality and indigenouslyness of the information; accessibility due to GL’s uniqueness and challenge in its inclusion in indexing and citation database; and content reliability and validity due to the often lack of peer-review process (Shrivastava, & Mahajan, 2021; Gul et al., 2021, Bickley, Kousa, & Thelwall, 2020). Hence, it is useful to discover and understand the influence of a locally published GL, thereby confirming GL's value in research and development.

The paper aimed to determine the impact of the technology developed by Engr. Alexis Belonio through the GL “Rice husk gas stove handbook”. Specifically, it aimed to determine the impact of GL by: 1) identifying the nationality of the authors who cited it; 2) determining the most common publishers of the citing work; 3) identifying the publication type of the citing works; 4) determining the number of citations of every citing works, and; 5) identifying the disciplines where the GL was being used and defused.

Methodology

The study included 155 literatures that cited the GL Rice Husk Gas Stove Handbook by Engr. Alexis Belonio. The data were identified and extracted from Google Scholar using Harzing’s Publish or Perish software. The search string used to retrieve the data was the complete title of the GL “Rice husk gas stove handbook”. To verify that the GL was cited in the retrieved literatures, the in-text citations and reference list of each title were checked manually.

Citation analysis was used to determine the influence of the GL. The analysis covers the nationalities and the collaborative activities of the authors who cited the GL; publisher; and publication type. To further see the influence, the analysis will also include the number of citations generated by the citing article and the discipline where the GL is prominently being used and being diffused. Frequency counts, percentages, and mean were used to describe the results.

Results

The GL written by Belonio was cited by 155 literatures which was written by 398 authors from 31 countries. Seventy-five (48.4%) of the 155 literatures were cited 797 times. This reflects the wide extent that the GL was used directly and indirectly.

Table 1 shows the number of authors per continent, where Asia leads with 274 authors (68.84 %), followed by Africa with 69 authors (17.34 %) and Europe with 8.79 % (35 authors).

Continent	No. of authors (f)	Percentage (%)
Asia	274	68.84
Africa	69	17.34
Europe	35	8.79
North America	15	3.77
South America	3	0.75
Australia	2	0.5
TOTAL	398	100

Table 1. Number of authors per continent

Moreover, the table below (Table 2) reflects the number of countries per continent where the nationality of the citing authors belongs. Asia included 10 countries (32.3%), followed by Africa and Europe, with eight (25.8%) countries each. North and South America consist of two (6.5%) countries each, and Australia only has one (3.2%).

Continent	No. of countries (f)	Percentage (%)
Asia	10	32.3
Africa	8	25.8
Europe	8	25.8
North America	2	6.5
South America	2	6.5
Australia	1	3.2
TOTAL	31	100

Table 2. Number of countries per continent

It can be assumed that since Asia dominates the rice production and consumption around the globe (Rice Alamac, 2013) and where more than 90% of rice is produced and consumed in the continent (CGIAR, 2019), many researchers notice the potential of the by-products of rice production.

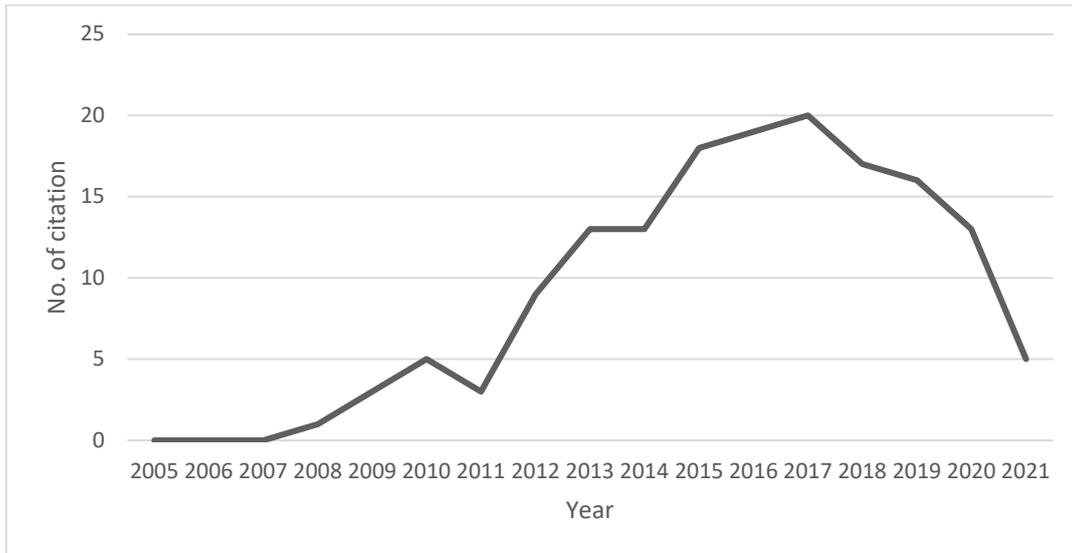


Figure 1. Number of citations per year

Since its publication in 2005, the GL has been cited 155 times. The work was cited at its highest in 2017, with 20 citations, as shown in Figure 1. On average, the work was cited 11 times a year in the course of 17 years. The peak was within 2015-2017, and in those years, Indonesia was the country with the highest number of citations. According to Faizah and Husaeni's (2018) work about the development of consumption and supplying energy in Indonesia's economy, the energy consumption from 2007 to 2017 fluctuated. Specifically, the energy consumption in households increased during the said years. The energy used by households includes biomass, gas, kerosene, LPG, and electricity. The types of biomass energy consumed by households are firewood, charcoal, and others used for cooking. Increased consumption in this sector is due to an increase in the number of family members and households in Indonesia. It can be assumed that this might be why researchers have ventured into researching the potential of rice husk as an alternative biomass energy source.

The result of the citation analysis also revealed that almost half of the first authors (49.03%) were Indonesian. It is also the same for the number of co-authors, where 98 (40.33%) of them are Indonesian out of 243 co-authors. The majority (90 %) of the first authors were affiliated with Academic Institutions, and the remaining proportions were connected with Government institutions (Figure 2).

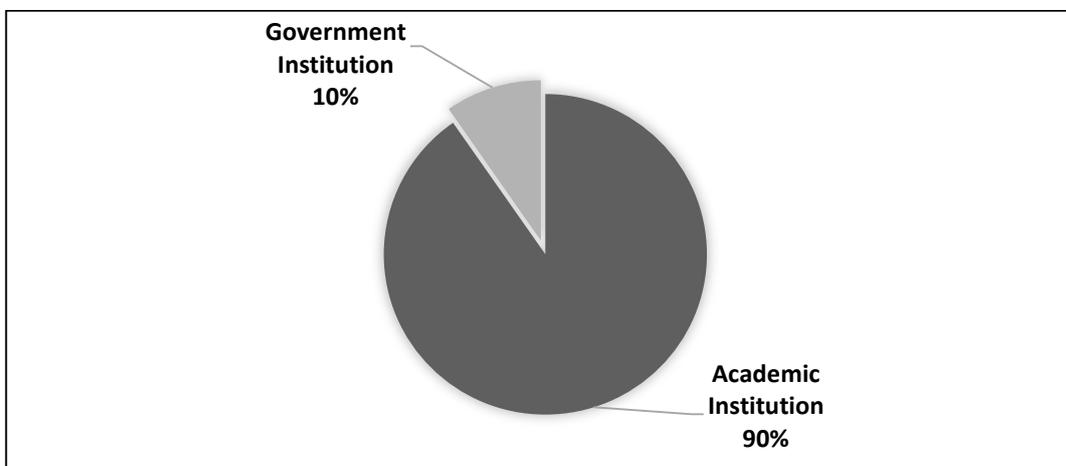


Figure 2. First author affiliation

Institution	f	%
Universitas Muhammadiyah Surakarta, Indonesia	19	11.45
Universitas Sriwijaya, Indonesia	5	3.01
Maharana Pratap University of Agriculture and Technology, India	4	2.41
Nong Lam University, Vietnam	4	2.41
Universitas Gadjah Mada, Indonesia	4	2.41
Universitas Jember, Indonesia	4	2.41
Universitas Sultan Ageng Tirtayasa (UNTIRTA), Indonesia	4	2.41
Bahir Dar Institute of Technology, Ethiopia	3	1.81
Indian Institute of Technology, India	3	1.81
McGill University, Canada	3	1.81

Table 3. Authors' affiliated institutions

Out of 101 institutions identified, the top 10 institutions with the greatest number of frequencies are all academic institutions (Table 3). Five of them come from Indonesia, two from India, and one from Vietnam, Ethiopia, and Canada, respectively. Universitas Muhammadiyah Surakarta, Indonesia ranks first with 11.4% followed by Universitas Sriwijaya, Indonesia with 3%.

Only a few (9%) of the citing literatures have a foreign collaboration in their work. 91% of the authors did their work without collaboration from other countries or nationalities. Most of the literatures that have been extracted were Journal Articles with 54.19% or 84 citing literatures. There were also Theses/Dissertations with 27.10% (42 citing literatures) followed by Conference Proceedings with 13.55% (21 citing literatures). Research Documents (including Research Articles, Projects, and Reports), Monographs, Technical Documents, and Bulletins were also identified and analyzed, as shown in Table 4. Moreover, most of the citing literatures (64%) are Open Access (Figure 3).

Type of Publication	f	%
Journal articles	84	54.19
Theses/Dissertations	42	27.1
Conference Proceedings	21	13.55
Research Documents	3	1.94
Monographs	2	1.29
Technical Documents	2	1.29
Bulletins	1	0.65
TOTAL	155	100

Table 4. Type of Publication of the citing literatures

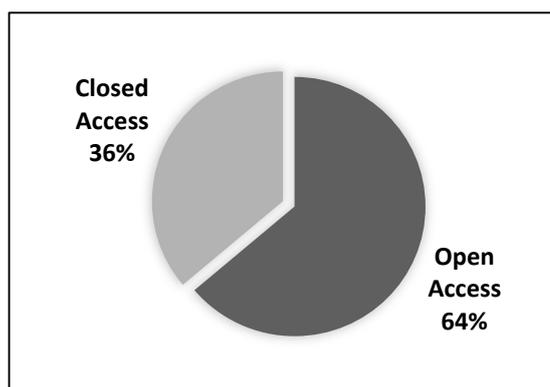


Figure 3. Type of access

The result of the analysis has identified 125 journals. Reflected in Table 5 are the journals with the most number of citing literature. Universitas Muhammadiyah Surakarta leads with 10.97%, followed by Renewable and Sustainable Energy Reviews (2.58 %) and Universitas Jember with 1.94%. Seven of the 10 (italicized) are Open Access journals.

Journal	f	%
Universitas Muhammadiyah Surakarta	17	10.97
Renewable and Sustainable Energy Reviews	4	2.58
<i>Universitas Jember</i>	3	1.94
<i>Berkala Fisika</i>	2	1.29
<i>Berkalan Ilmiah Teknologi Pertanian</i>	2	1.29
EPD Congress	2	1.29
<i>International Letters of Chemistry, Physics and Astronomy</i>	2	1.29
<i>IOSR Journal of Applied Chemistry (IOSR-JAC)</i>	2	1.29
<i>Journal of Sustainable Bioenergy Systems</i>	2	1.29
<i>Jurnal Teknik Kimia Indonesia</i>	2	1.29

Tables 5. Journals with the most number of citing literature

Publisher	f	%
Universitas Muhammadiyah Surakarta	15	9.6
<i>Springer</i>	9	5.8
<i>Elsevier</i>	6	3.8
IOP Publishing	4	2.6
<i>Taylor & Francis</i>	4	2.6
UMS Institutional Repository	3	1.9
Universitas Jember	3	1.9
Universitas Sriwijaya	3	1.9
Digital Repository Universitas Jember	2	1.3
FLUIDAS Asociația Natională Profesională De Hidraulică Si Pneumatică	2	1.3

Table 6. Publishers with most number of citing literature published

Out of 106 publishers identified, Universitas Muhammadiyah Surakarta is the highest with 9.6%. Springer follows with 5.8 %, and Elsevier falls third with 3.8%. Table 6 shows the publishers with the most number of published citing literature. Only three commercial publishers (italicized) are included on the publishers with the most number of published citing literature, which includes Springer, Elsevier, and Taylor & Francis.

The GL was mainly used in the field of Mechanical Engineering (40.65%) followed by Agricultural Engineering (39.35%) and Industrial Engineering (11.61%), as shown in Table 7. It can be assumed that Mechanical Engineering ranked first because authors or researchers are using Belonio’s work to create their own modified or specific version of the invention. Also, they referred to the GL for more information regarding the use of rice husk as biomass energy.

Discipline	f	%
Mechanical Engineering	63	40.65
Agricultural Engineering	61	39.35
Industrial Engineering	18	11.61
Agriculture	11	7.10
Chemical Engineering	1	0.65
Electrical Engineering	1	0.65
TOTAL	155	100

Table 7. Disciplines where the GL was diffused

governments to utilize the potential of rice husks as an alternative and renewable biomass energy (Quispe, Navia, & Khhat, 2016; Kate & Chaurasia, 2018).

We can assume that majority of the authors that cited the GL were from Asia because of the potential and abundance of rice husks, an agricultural waste that can be utilized for alternative biomass energy since the region is also leading in rice production (GRiSP, 2013). This is supported by Pode, Diouf, & Pode's (2015) statement that biomass energy was being developed and widely used in Asian countries. Moreover, Quispe, Navia, and Kahhat (2017) discussed in their paper the energy potential from rice husk per year per region, where Asia leads, followed by America, Africa, and Europe.

The majority of the literature was open access journal articles, mostly published in academic institutions through their repositories. This result is in line with the statement of Gul et al. (2020) in their paper that the role of Open Access Repositories (OAR) in disseminating grey literatures are evident. Grey literature continues to play an essential part in the research endeavor of researchers (Palcullo, Geromiano, & Superio, 2021). Predominantly, repositories are institutional, disciplinary, aggregating, or governmental. Specifically, Institutional Repositories (IR) are managed by academic or research institutions (Gul et al., 2021). Tsunoda, Sun, Nishizawa, & Liu (2017) stated that IRs "aim to provide open access to institutional research output, to create global visibility for institutions' research, and to store and preserve other institutional digital assets, including unpublished or otherwise easily lost grey literature such as theses, working papers or technical reports."

Conclusions and Recommendations

This study exhibited that the role of Grey Literature as a source of information stays essential, especially for researchers. The top users of the GL came from the countries which are also considered leading countries in rice production and consumption. The GL was used in different disciplines, mainly in Mechanical Engineering and Agricultural Engineering, and most of the literature that used the GL was available as Open Access.

The study shows that even locally-published grey literature could be of value to the international community; thus, publishers must ensure that the distribution of these resources must not be limited to the local community only.

Furthermore, Institutional repositories play important roles in disseminating and preserving these publications; thus, whenever possible, the establishment of such is recommended.

References

- Agoot, L. (2020, December 22). *PH logs highest rice production rate at 19.44M metric tons: DA*. Philippine News Agency. <https://www.pna.gov.ph/articles/1125473>
- Belonio, *Rice husk gas stove handbook*. Central Philippine University.
- 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.
- CGIAR. (2019). International Rice Market/Trade. Retrieved from Ricepedia - CGIAR: <http://ricepedia.org/rice-as-commodity/international-rice-market-trade>
- Faizah, S. T., & Husaeni, U. A. (2018). Development of consumption and supplying energy in Indonesia's economy. *International Journal of Energy Economics and Policy*, 8(6), 313-321.
- Fung, V. W. H., & Jenkins, B. M. (2003). Biomass power development for the Philippines. In *2003 ASAE Annual Meeting* (p. 1). American Society of Agricultural and Biological Engineers.
- GRiSP (Global Rice Science Partnership). (2013). *Rice almanac* (4th ed.). International Rice Research Institute.
- Gul, S., Shah, T. A., Ahmad, S., Gulzar, F., & Shabir, T. (2021). Is grey literature really grey or a hidden glory to showcase the sleeping beauty. *Collection and Curation*, 40/3, 100-111.

- Kate, G. U., & Chaurasia, A. S. (2018). Gasification of rice husk in two-stage gasifier to produce syngas, silica and activated carbon. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 40(4), 466-471.
- Kumar, S., Sangwan, P., Dhankhar, R. M. V., & Bidra, S. (2013). Utilization of rice husk and their ash: A Review. *Research Journal of Chemical and Environmental Sciences*, 1(5), 126-129.
- Lim, J. S., Manan, Z. A., Alwi, S. R. W., & Hashim, H. (2012). A review on utilisation of biomass from rice industry as a source of renewable energy. *Renewable and Sustainable Energy Reviews*, 16, 3084-3094.
<https://doi.org/10.1016/j.rser.2012.02.051>
- Mason, M. K. (2009). Grey Literature: History, Definition, Acquisition and Cataloguing. Retrieved 27 Oct. 21 from <http://www.moyak.com/papers/grey-technical-literature.html>
- Mofijur, M., Mahlia, T. M. I., Logeswaran, J., Anwar, M., Silitonga, A. S., Rahman, S. M. A., & Shamsuddin, A. H. (2019). Potential of Rice Industry Biomass as a Renewable Energy Source. *Energies*, 12(21), 4116.
doi:10.3390/en12214116
- OECD/FAO. (2021). *OECD-FAO Agricultural Outlook 2021-2030*. Paris: OECD Publishing.
<https://doi.org/10.1787/19428846-en>.
- Palcullo, V. E. V., Geromiano, J. F., & Superio, D. L. (2021). Grey literature usage among Filipino aquaculture researchers: A bibliometric analysis of research from 2009 to 2018. In T. Vicary & M. Kalentsits (Eds.), *Enabling Grey Literature Discovery to Benefit Aquatic Science, Fisheries and Aquaculture, Terengganu, Malaysia, 25 September, 2019* (pp. 17–19). Rome, Italy: Food and Agriculture Organization (FAO).
- Pode, R., Diouf, B., & Pode, G. (2015). Sustainable rural electrification using rice husk biomass energy: A case study of Cambodia. *Renewable and Sustainable Energy Reviews*, 44, 530-542.
- Punin, W. (2020). Evaluation of the thermal efficiency and a cost analysis of a new rice husk gas cookstove for the rural areas of Northern Thailand. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 1-13.
- Quispe, I., Navia, R., & Kahhat, R. (2017). Energy potential from rice husk through direct combustion and fast pyrolysis: a review. *Waste management*, 59, 200-210.
- Schöpfel, J., & Farace, D. J. (2009). Grey literature. In *Encyclopedia of library and information sciences* (pp. 2029-2039). CRC Press.
- 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.
- Simeon, L. M. (2016, November 27). Government promotes rice husk as alternative energy source. *The Philippine Star*. <https://www.philstar.com/business/agriculture/2016/11/27/1645845/government-promotes-rice-husk-alternative-energy-source>
- Strauss, G. (2016, October 20). *This Cooker Uses Rice Husks as a Cheap, Green Fuel Source*. National Geographic, Explorer Moments. <https://www.nationalgeographic.com/science/article/alexis-belonio-explorer-moments-food-innovation>
- Tillett, S., & Newbold, E. (2006). Grey literature at The British Library: Revealing a hidden resource. *Interlending & Document Supply*, 34(2), 70-73.
- Tsunoda, H., Sun, Y., Nishizawa, M. and Liu, X. (2017, August 16-17). *Grey literature in institutional repositories: a case study on the world top 100 universities* [Paper presentation]. IFLA WLIC 2017 WRO CLAW: Satellite Meeting: Serials and Other Continuing Resources Section & Acquisition and Collection Development Open Access: Action Required, European Solidarity Center (ESC), Gdańsk.
- Vinluan, Jr., F. D. (2002). Carbonized rice husk (CRH): Production, export potential and its applications for agriculture & industry [Abstract]. *Philippine Agricultural Mechanization Bulletin* 9(2).
<http://scinet.dost.gov.ph/union/ShowSearchResult.php?s=2&f=&p=&x=&page=&sid=1&id=Carbonized+rice+husk+%28CRH%29&Mtype=ANALYTICS>

Czech National Repository of Grey Literature

The logo for NUSL (National Library of Technology) features the letters 'NU' in white and 'SL' in white with a green vertical bar to its right, all on a black 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

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

NUSL
czech
national
repository
of grey
literature

Grey literature and the role of women pioneers in Qajar era (1910-1920)

Somayeh Sadat Hashemi, Roya Aminalroaya, and Nooshin Hakhamaneshi,
National Library and Archives of Iran

Abstract

Iranian society in the Qajar era, despite having an ancient civilization and culture, had a medieval society with national and religious prejudices. During this era, due to the society's view of women and the lack of importance to their prosperity, their education was not considered. After constitutional revolution, newspapers and magazines became more active and their number increased day by day. In this time, newspapers became very popular among the people, and their news covered not only political and social subjects, but also the news of women of other nations in the form of apparent values, duties and status in the family. This news and information and intellectual progress led women to cooperate with the press of that time; because they considered it is the perfect tool to express their thoughts with others and connecting abroad. In this regard, women started publishing magazines in order to awaken and increase the awareness of women in the society in which the first magazine was published by women was "Danesh".

In this research, the articles of women's journals, that were published during 1910-1920, studied in order to obtain an understanding of the information needs of women in that period. This study surveys the topics and the number of articles written or translated by women. The research population consisted of 116 issues from six titles of the women journals and a total of 569 articles were reviewed.

A review of these journals shows that women initially tried to cover subjects such as childbearing, marriage, health, etc. in their own media. The purpose of these journals was to take women out of the space created by their dominant culture or society, to bring them into the community, and to make them aware of the right to equality with men, while preserving their dignity as mothers and females. These media were a window to address the most basic information needs of women, such as health and psychological topics. It was especially important to inform women about the health and protection of children at a time when infectious diseases such as smallpox were pandemic.

Keywords: Iranian women, Qajar era, Grey literature, Women Journals

Qajar dynasty

The reign of the Qajar dynasty in Iran (1796 – 1925) is recognized as a period of dynamic political, economic, and cultural transformations. Founded in the late 1700s, the Qajars ruled for the next century and a half. The last decades of the dynasty's rule were also marked by Iran's first twentieth-century revolution, the Constitutional Revolution of 1906, which resulted in the formation of an elected parliament and the drafting of a national constitution, both of which still comprise the backbone of Iran's government¹.

Social status of women during the Qajar

"Behind the closed doors at home, prohibited from everything in life, education, training and social life, women are regarded as mindless, like infants; they are confined to the burdens of household work and childbearing and are considered the slaves and servants of their husbands," wrote Bibi Khanum Astarabadi (1852–1920), an outspoken and prominent Qajar woman. Similarly, in describing women's absence in public, Mohammad Ali Jamalzadeh, a noted novelist commented: "No women can be seen in this country of men, but strangely, half of the walking population in the streets is wrapped in black bags from head to the toe without even an opening to breathe."

Yet, these invisible women were capable of action, as in the Constitutional Revolution of 1905–1911 when many women gathered in the streets of Tehran took off their veils and shouted: “Long live freedom. . . . We must . . . live the way we want!”

Women were primarily confined to the household and reproduction. Their three-piece dress consisting of the chador (a long veil that covered them from head to toe), the rubandeh (a short veil that masked the face), and the chaqchur (very loose trousers) that signified their separate world; it assured them space and identity as *zai’feh*, or the weak sex and status as *moti’eh*, or those obedient to men’s will.

Women spent most of their lives in the private world of the family. Indeed, a common name for a wife was *manzel* (the home). Rich or poor, women were confined to, and devoted their entire lives to the family.

Focusing on women’s journals, this historical research discusses women’s experiences in the family and their work, religion, and politics at the turn of the twentieth-century².

In sum, despite variations in different classes, women were primarily confined to the private and secluded world of the family. Patriarchal power also varied by class. The higher women were on the social scale, the more secluded and less mobile they were. By contrast, less-privileged women were more mobile and less secluded. Class and patriarchy acted together to shape women’s lives; together, they affected women’s work both within and outside the family.

Depending on the kinds of work they performed, working Iranian women often combined child rearing with their tasks in the larger economy. Not all women worked outside the family. Those who did, they struggled against poverty, whether single or married, or as child laborers. While these women encountered male domination both in the household and in the marketplace, the economically secure women stayed home and experienced patriarchy more directly there. Gender subordination varied by class insofar as women’s work was concerned. Women’s work was diverse. Many poor women worked as carpet weavers, vendors, domestic laborers, and seamstresses³.

The beginning of women's activity as a journalist and the process of content production for women in different classes of society

With the Constitutional Revolution and the beginning of the Enlightenment, we see that the publication of journals by women for women is not far from the beginning of the publication of newspapers (in general) in Iran.

The director and editor in chief of women's periodicals were mostly women, and each of them had significant social activities (women's movement activist, writer, doctor, etc.) and was fluent in one of the foreign languages.

The women’s journal of this era was: “*Danesh*”, “*Shokoufeh*”, “*Zaban-e Zanan*”, “*Name-ye Banovan*”, “*Jahan-e Zanan*” and “*A’alam-e Nesvan*”⁴.

Danesh Journal

Danesh was the first women's magazine; the first issue was published in 1910, four years after the Constitutional Revolution. The director was *Dr. Kahal*. The contents of this magazine were in moral subjects, housekeeping, childbearing, and marriage didn’t talk about the politics. Its goal was to educate women and girls and to teach morals to women. Although it was the first experience of women in journalism, it succeeded to a large extent in meeting the goals set out at the outset and took effective steps to introduce women to society⁵.

Shokoufeh Journal

The first issue of the *Shokoufeh* was published in 1913 under the management of *Maryam Amid*. Beautiful and instructive cartoons were printed in each issue of *Shokoofeh* and critically reviewed the superstitions and old customs of the society. Most of *Shokoofeh*'s articles were related to women's schools, their exams, and other matters related to women and girls. It also covered news and articles about women around the world⁶.

Zaban-e Zanan Newspaper

In 1919, *Sedigheh Dolatabadi* (director) published the *Zaban-e Zanan* newspaper in Isfahan. This is the first women's newspaper to have the word 'woman' (*Zanan*) in its title. During its publication, the newspaper advocated for freedom and the development of culture among women in the country⁷.

Jahan-e Zanan Journal

It was a special magazine on women's rights and education that *Fakhr Aafagh Parsa* started publishing as a director in Mashhad in 1920 and later continued to publish in Tehran. The content of this Magazine was more about European-style women's liberation and unveiling and equal rights⁸.

Name- ye Banovan Newspaper

Its editor in chief was *Shahnaz Azad*, whose first issue was published in 1920 in Tehran. The pro-unveiling newspaper also published articles written exclusively by women, as well as national and international news⁹.

A'alam-e Nesvan Journal

Its editor in chief was *Nawabeh Safavid*, and established in 1920 in Tehran. The articles were written by the American School graduates for Girls in the Presbyterian missionary compound in Tehran. The subject of the articles was extensive and included medical reports, housekeeping tricks, fashion in the West, literary works, and news from feminist movements around the world. The magazine also published articles denouncing the early marriage of girls, the lack of women's political rights in Iran, and the *hijab* (=veil)¹⁰.

Findings

A review of these journals shows that women initially tried to cover subjects such as childbearing, marriage, health, etc. in their own media. These media were a window to address the most basic information needs of women, such as health and psychological topics. It was especially important to inform women about the health and protection of children at a time when infectious diseases such as smallpox were pandemic. Later, the purpose of these journals changed to take women out of the space created by their dominant culture or society, to bring them into the community, and to make them aware of the right to equality with men, while preserving their dignity as mothers and females.

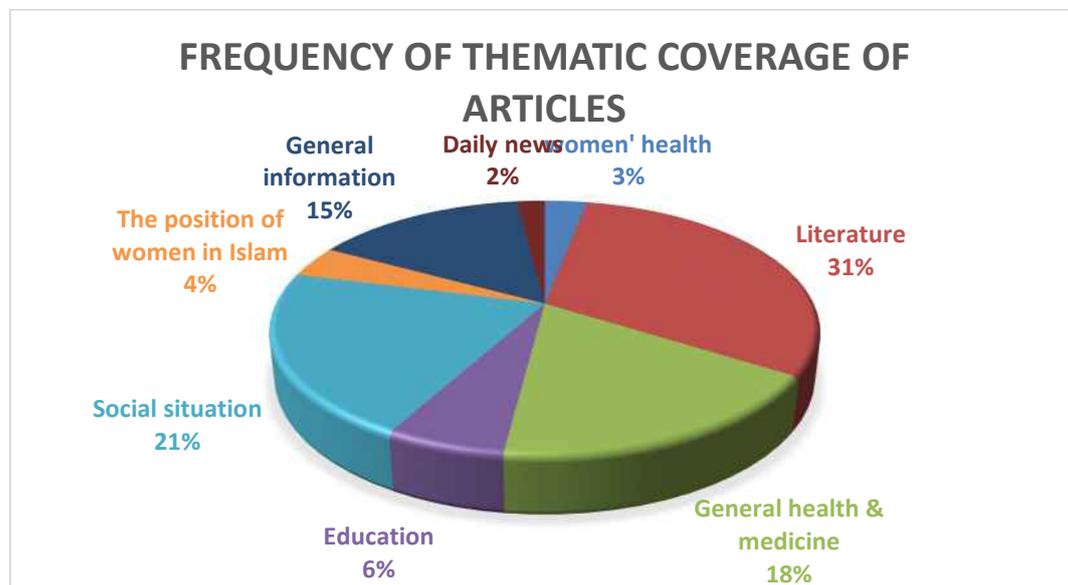


Chart1. Frequency of thematic coverage of articles

Out of 569 articles, the most thematic coverage is in the fields of literature with 31% and women’s life style and social situation with 21%. Also, the least covered topic is the daily news about political issues (2%). The publication of poems on the subject of women's status and dignity, as well as stories and myths in which women have played a role, are other points to consider in these articles.

Due to the awareness of women in the years since the publication of newspapers, these publications also gradually addressed issues such as women's rights and their role in society, which, of course, was not very pleasant to the society of that day and led to their closure. This is a reason for the small number of articles in the field of current issues and political news.

As Figure 2 shows, most articles are written by women (55.54%). The lowest number of articles was written by men (12.3%); Also, the highest number of translations with 8.26% was done by women and the lowest by 1.05% by men.

One of the reasons is the prejudice of some publications in publishing articles written by women. As some of them like *Zaban-e Zanan* stated in their editorial, they would not publish at all if they received an article from the gentlemen: “Only the letters and articles of women and girls are accepted. Men should not bother to write something that will not be accepted”.

The unfamiliarity of many women writers with a second language caused the number of translated articles to differ significantly from other articles. It should be noted that this is a positive point and encouraged women to write.

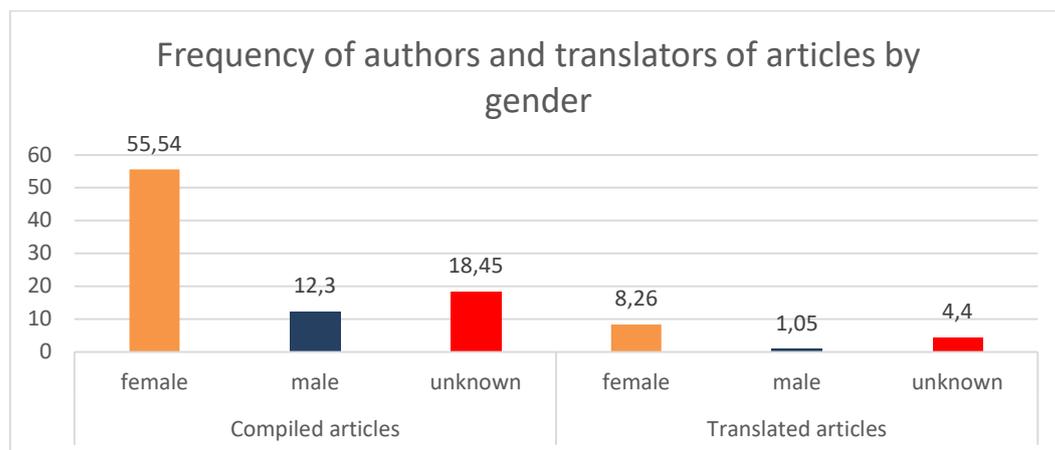


Chart2. Frequency of authors and translators of articles by gender

Although the publication of journals for women caused significant changes, public opinion still did not accept the presence of women in society. Therefore, many women published their articles and translations in journals under abbreviated or anonymous names, and the author of 162 articles (28.47%) is not known to have been a woman or a man.

Conclusion

The articles written in the specialized journals of women in the Qajar period were mostly about topics related to home, family and children, and the daily news (included the social and economic news). There are no political subjects in the initial issues of these articles. Perhaps the reason for this is the lack of connection with other countries and societies, which caused them to stay away from new information and news and promote superstition.

Gradually, as these publications found their place in the society, a number of them, such as *Zaban-e Zanan* and *A’alam –e Nesvan*, addressed topics related to women's political rights and current political issues.

In this regard, in order to achieve individuality, women had to achieve the following components: the right to choose and freedom to decide, accept responsibility for choice, rely on rationality with the tools of knowledge and science, desire for dynamism and

progress in life, attention to individual rights and social. Permission to publish articles for women shows the awakening of society and the need for change in life. The gradual launch of journals founded and run by knowledgeable and educated women is a revolution in social life and a change in women's attitudes.

As these journals show in subsequent years and the content of their articles, women were no longer indifferent to internal and external events at this stage as in the past. By studying newspapers, participating in association meetings, literacy, etc., they realized the difference between the conditions of their country and other societies and called for the improvement of their situation. Women also realized the differences between societies and called for an end to this discrimination.

References

- ¹. Women's Worlds in Qajar Iran: A Digital Archive and Website. Retrieved in 2021, Sep. from https://www.neh.gov/sites/default/files/inlinefiles/harvard_univerty_womens_worlds_in_qajar_iran_a_digital_archive_and_website.pdf
- ². https://doi.org/10.1017/CCBOa9m78b0r5id11g5e1_0B3o80o.k00s3
- ³. McElrone, Susynne M. (2005). Nineteenth-Century Qajar Women in the Public Sphere: An Alternative Historical and Historiographical Reading of the Roots of Iranian Women's Activism. *Comparative Studies of South Asia, Africa and the Middle East*. V. 25, No. 2. pp. 297-317
- ⁴. Babran, Sedigheh (2002). Women's special publications (in Persian: Nashriyat -e vije Zanan). Tehran: Roshangaran
- ⁵. Danesh Journal (1910)
- ⁶. Shokoufeh Journal (1913)
- ⁷. Zaban-e Zanan Newspaper (1919)
- ⁸. Jahan-e Zanan Journal (1920)
- ⁹. Name- ye Banovan Newspaper (1920)
- ¹⁰. A'alam-e Nesvan Journal (1920)



GreyGuide: an example of Open Access Publishing in GL

Stefania Biagioni and Carlo Carlesi, CNR-ISTI-Pisa Italy
Dominic Farace, GreyNet International, Amsterdam, NL

Repository and Portal to Good Practices and Resources in Grey Literature

Welcome to the GreyGuide: point of access to Grey Literature and Open Access Resources <http://greyguide.isti.cnr.it/>

PARTNERS



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



CNR ISTI
NeMis Lab

<http://www.isti.cnr.it>

CONTENT PROVIDER

Dominic Farace
GreyNet
International
Netherlands

SYSTEM MANAGER

Stefania Biagioni
CNR-ISTI
NeMIS Lab
Italy

TECHNICAL DEVELOPER

Carlo Carlesi
CNR-ISTI
NeMIS Lab
Italy

COLLECTIONS

BIO: Who is in
Grey Literature

GLA: Conference
Proposals

GLP: Conference
Papers

RGL: Resources
in Grey
Literature

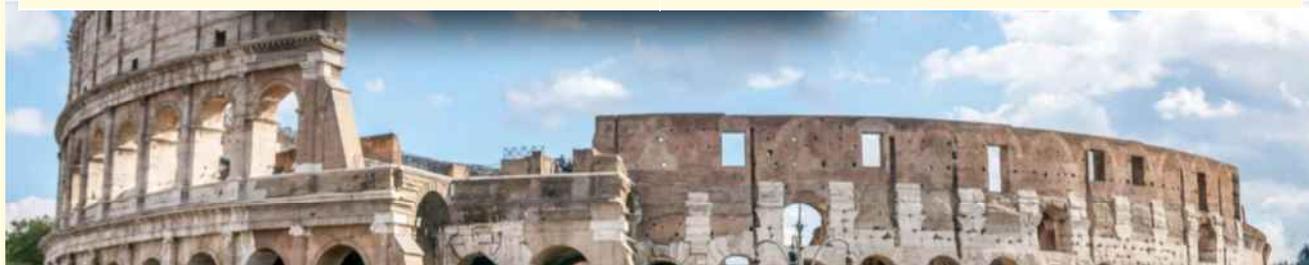
The GreyGuide is
steered by
GreyNet's Resource
Policy Committee
(RPC)



Remember
to endorse
The Pisa
Declaration

The poster shows the goals achieved in the last 5 years, the progress, new features and new resources made available by GreyGuide in support of Open Access Publishing. In 2015, GreyNet International carried out an online survey among its stakeholders in order to determine their use of its sustained information resources.

Now five years on, having benefited from technical developments, the migration of hundreds of metadata full-text records, and the addition of enriched fields and functionality, the GreyGuide offers GreyNet a testbed from which to map and measure its capacity in open access publishing. The population of this study is drawn from digital resources accessible via both the GreyGuide Portal and Repository.



Scenario

GreyNet's web-access portal and repository is the GreyGuide – an internet resource that is fully open access compliant, launched in 2013 as a collaborative effort between GreyNet International and CNR-ISTI, NeMIS Lab, Pisa, Italy. GreyGuideRep is a platform supporting document submission, curation, preservation and sharing.

Objective

Pays particular attention to Open Access Publishing . Shares Research and Knowledge in the field of Grey Literature via the GreyGuide Portal and Repository. Meets the needs of different levels of users and increases the visibility and reuse of documents and research data.

What's on

- **New resources** in Document Share enabling wider public access to Grey Literature;
- **DOIs** for GL-Conference Papers and diverse types of RGL documents;
- **More accredited Identifiers:** OpenDoar, DOI, ORCID, CC BY;
- **Open access to the largest Collection** of GL Conference Posters and Slides.
- **We are Joining OpenAire**

The way in which the digital resources are openly accessible



Web Access Portal (Document Share)

Conference Posters	Conference Slides	Program Books	Conference Proceedings	GreyNet Newsletters	GL Advertorials	Grey Forum Series	GL Conference Videos
------------------------------------	-----------------------------------	-------------------------------	--	-------------------------------------	---------------------------------	-----------------------------------	--------------------------------------

Portal Jump Page

Conference Preprints	OpenAIRE GreyNet Publications	Research Datasets	GreySource Index	GL Guides	GL TIB-AV Videos
--------------------------------------	---	-----------------------------------	----------------------------------	---------------------------	----------------------------------

Web Access Repository

GLA Conference Abstracts	GLP Conference Papers	BIO Biographical Notes	RGL Resources of GL
--	---------------------------------------	--	-------------------------------------

GreyGuide is a point of access to other Grey Literature Resources

- WorldWideScience.org Gateway
- INIS, International Nuclear Information System Repository
- NUSL, National Repository of Grey Literature
- TIB AV-Portal, A web-based platform for quality-tested scientific videos
- e-LIS Repository
- APO, Analysis & Policy Observatory

GreyGuide Repository Accredited Identifiers



Data from “Exploring Next Generation Grey”

<https://doi.org/10.17026/dans-xrg-2gf6>

Joachim Schöpfel, University of Lille, France

<https://orcid.org/0000-0002-4000-807X>

<https://ror.org/02kzqn938>

Project Moderator / Text Writer

Dominic Farace, GreyNet International, Netherlands

<https://orcid.org/0000-0003-2561-3631>

<https://ror.org/01pxfxj80>

Project Administration / Text Reviewer

David Baxter, University of Alberta, Canada

<https://orcid.org/0000-0001-5235-6728>

<https://ror.org/0160cpw27>

Questionnaire

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

<https://orcid.org/0000-0001-7323-3786>

<https://ror.org/05kacka20>

Questionnaire

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

<https://orcid.org/0000-0002-4793-3727>

<https://ror.org/05kacka20>

Questionnaire

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

<https://ror.org/031q21x57>

Questionnaire

Veronika Potočnik, National and University Library, Slovenia

<https://ror.org/04tjx1w16>

Questionnaire

Dobrica Savić, Nuclear Information Section, NIS-IAEA, United Nations

<https://orcid.org/0000-0003-1123-9693>

<https://ror.org/00gtfax65>

Questionnaire

Abstract

The GL2021 Conference¹ offered the many and diverse communities of practice in the field of grey literature a unique opportunity to collaborate in addressing and defining the next phase in the digital transformation of grey literature. In preparation for this conference, a panel session on the future of grey literature was planned on the program; and, in advance, an online survey was carried out among GreyNet’s own community of practice in the field of grey literature.

Keywords

Grey literature; scientific and technical information; libraries; archives; museums; advocacy

Subject Area

Library and information sciences

Methods Applied

• Steps

A selection of five panel members was made among the GreyNet’s community of practice representing different subject areas and fields of interest. Each of the five invited panelists was asked to provide a topic they consider of significant importance for grey literature. The Panel Moderator checked that there were no duplications among the five topics. Once the topics had been decided, each of the five panelists was then asked to submit two questions pertaining to their topic. This then accounted for the 10 questions in the survey. In order to standardize responses to the 10 questions, the choices of response were limited to: Strongly Agree Agree Uncertain Disagree Strongly Disagree

Each response to a question allowed for further comment. The online questionnaire was then launched on the SurveyMonkey platform².

• Sampling strategy

The link to the online survey was made openly accessible via GreyNet’s Distribution List (890 recipients) and social media: Facebook³ (158 friends), LinkedIn⁴ (672 members), and Twitter⁵ (1201 followers). While the population of the survey was not controlled, it is considered that all of the potential respondents have some level of affiliation with grey literature. Two reminders were sent out before the close of the survey.

Survey Population	Number of Survey Respondents	Percentage of Questions Answered	Average Number of Comments made per Question
Uncontrolled	40	99.97%	8.2

• Quality Control

The survey questions were double-checked by the panel members and the panel moderator. There was no specific control carried out on the data acquired from the survey. None of the categories of responses were grouped or otherwise normalized. All of the recorded comments were in line with the questioning, which may allow one to assume that there is no cause to question the validity of the responses.

Dataset Description

File name:	GL2021 Survey Results
Format:	PDF
Size:	308 KB
Creation dates:	from 2021-06-20 to 2021-09-14
Language:	English
License:	CC0 Waiver - no rights reserved
Archive name:	DANS EASY Archive
Publication date:	2021-11-22
DOI:	10.17026/dans-xrg-2gf6
URN:	urn:nbn:nl:ui:13-f3-av5w

Potential Reuse of the Data

The results of the survey could be used in a Position Paper, which may include the formal statements presented by the five panellists⁶. The data can also be useful for other researchers and information professionals interested in the development of scientific and technical information in general and grey literature in particular. Its reuse could be helpful in addressing issues dealing with documentary and cultural heritage, digital preservation, citizen science, peer review, as well as legal and policy matters. On a more critical note, the data is limited by the number of respondents to the survey – 40 in total. It remains a fact that the number of respondents does not formally allow for the expression of results in percentages. The data however remains preserved in a national archive⁷, which carries the CoreTrustSeal⁸ and by way of this data paper demonstrates compliance with FAIR data principles⁹.

Linked References

¹ <https://www.youtube.com/channel/UCVaYldfpwZoQwAK2Dsqu-wQ>

² <https://www.surveymonkey.com/>

³ <https://www.facebook.com/greynetinternational>

⁴ <https://www.linkedin.com/groups/3718857/>

⁵ <https://twitter.com/GreyLitNet>

⁶ https://www.youtube.com/playlist?list=PL38n__kpNAudqSttKuvNOPU6MAH7IYWMS

⁷ <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:68541>

⁸ <https://www.coretrustseal.org/wp-content/uploads/2018/04/DANS-Electronic-Archiving-SYstem-EASY-.pdf>

⁹ <https://www.force11.org/group/fairgroup/fairprinciples>

Appendix: Exploring Next Generation Grey – Panel Survey Results

Q1

Grey literature is part of the everyday working routine of librarians and information professionals operating in non-academic institutions, for example museums, archives and public libraries.

- Answered: 39
- Skipped: 1

ANSWER CHOICES-	RESPONSES-
- Strongly agree	41.03% 16
- Agree	41.03% 16
- Neither agree nor disagree	5.13% 2
- Disagree	10.26% 4
- Strongly disagree	2.56% 1
Total Respondents: 39	

Q2

In academic institutions, the affirmation of open science and open access principles significantly favor the production, publication, and retrieval of grey literature.

- Answered: 39
- Skipped: 1

ANSWER CHOICES-	RESPONSES-
- Strongly agree	28.21% 11
- Agree	38.46% 15
- Neither agree nor disagree	20.51% 8
- Disagree	7.69% 3
- Strongly disagree	5.13% 2
Total Respondents: 39	

Appendix: Exploring Next Generation Grey – Panel Survey Results

Q3

It will be difficult to find material in the library, whose process and handling is predominantly dependent on library resources.

- Answered: 38
- Skipped: 2

ANSWER CHOICES-	RESPONSES-
- Strongly agree	18.42% 7
- Agree	28.95% 11
- Neither agree nor disagree	36.84% 14
- Disagree	10.53% 4
- Strongly disagree	5.26% 2
Total Respondents: 38	

Q4

Small prints are an important part of each library and represents a fund of national cultural heritage.

- Answered: 38
- Skipped: 2

ANSWER CHOICES-	RESPONSES-
- Strongly Agree	26.32% 10
- Agree	47.37% 18
- Neither agree nor disagree	23.68% 9
- Disagree	0.00% 0
- Strongly Disagree	2.63% 1
Total Respondents: 38	

Appendix: Exploring Next Generation Grey – Panel Survey Results

Q5

The majority of future academic, scientific, and technical literature will be grey.

- Answered: 39
- Skipped: 1

ANSWER CHOICES-	RESPONSES-
- Strongly agree	17.95% 7
- Agree	20.51% 8
- Neither agree nor disagree	28.21% 11
- Disagree	23.08% 9
- Strongly disagree	10.26% 4
Total Respondents: 39	

Q6

The management of future grey and other literature, information, and data will converge.

- Answered: 39
- Skipped: 1

ANSWER CHOICES-	RESPONSES-
- Strongly agree	17.95% 7
- Agree	56.41% 22
- Neither agree nor disagree	20.51% 8
- Disagree	7.69% 3
- Strongly disagree	0.00% 0
Total Respondents: 39	

Appendix: Exploring Next Generation Grey – Panel Survey Results

Q7

Incentives for academics to participate in research published as grey literature must be improved.

- Answered: 39
- Skipped: 1

ANSWER CHOICES-	RESPONSES-
- Strongly agree	61.54% 24
- Agree	17.95% 7
- Neither agree nor disagree	15.38% 6
- Disagree	2.56% 1
- Strongly disagree	2.56% 1
Total Respondents: 39	

Q8

Peer-review processes are the most important method for raising the profile of research published as grey literature.

- Answered: 39
- Skipped: 1

ANSWER CHOICES-	RESPONSES-
- Strongly agree	35.90% 14
- Agree	28.21% 11
- Neither agree nor disagree	25.64% 10
- Disagree	7.69% 3
- Strongly disagree	7.69% 3
Total Respondents: 39	

Appendix: Exploring Next Generation Grey – Panel Survey Results

Q9

While often country specific, the legal and policy infrastructure is adequate in terms of funding, legal understandings, and policy protocols that support grey literature.

- Answered: 39
- Skipped: 1

ANSWER CHOICES-	RESPONSES-
- Strongly agree	0.00% 0
- Agree	15.38% 6
- Neither agree nor disagree	33.33% 13
- Disagree	43.59% 17
- Strongly disagree	7.69% 3
Total Respondents: 39	

Q10

The grey literature community should organize or develop an advocacy group that would actively promote its concerns to policy and decision makers.

- Answered: 40
- Skipped: 0

ANSWER CHOICES-	RESPONSES-
- Strongly agree	47.50% 19
- Agree	37.50% 15
- Neither agree nor disagree	12.50% 5
- Disagree	2.50% 1
- Strongly disagree	0.00% 0
Total Respondents: 40	

List of Participating Organizations

Amsterdam Public Library, OBA	Netherlands
Brigham Young University	USA
Central Philippine University	Philippines
Centre for Media and Celebrity Studies	USA
Coherent Digital	France
Data Archiving & Networked Services, DANS-KNAW	Netherlands
EBSCO Publishing	USA
EBSCO BeNeLux	Netherlands
GERiCO laboratory	France
German National Library of Science and Technology, TIB	Germany
GreyNet International	Netherlands
Information Today, Inc.	USA
Infrastructures for Science Laboratory, InfraScience	Italy
Institute of Information Science and Technologies, ISTI-CNR	Italy
International Council for Scientific and Technical Information, ICSTI	France
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 and Archives	Iran
National Library of Technology, NTK	Czech Republic
Nuclear Information Section, NIS-IAEA	UN/Austria
Saint Petersburg State University	Russia
Slovak Centre of Scientific and Technical Information	Slovakia
Social Science Research Network, SSRN-Elsevier	USA
Southeast Asian Fisheries Development Center	Philippines
TextRelease	Netherlands
University of Alberta	Canada
University of California, Irvine	USA
University of Florida; George A. Smathers Libraries	USA
University of Florida; UF Research	USA
University of Denver	USA
University of Lille	France
University of Rennes	France
University of Wisconsin—Milwaukee	USA
U.S. Department of Energy; Office of Scientific and Technical Information	USA
WorldWideScience Alliance	USA

Author Information

Baars, Chris

45

Supervisor Digital Services at DANS: project lead DANS Archives, NARCIS (National Academic Research and Collaborations Information System) and DataverseNL. Member of ORCID Trust Working Group and EuroCRIS. Specialties: Digital Preservation; Research Data Management; Reuse of Data; Metadata; Persistent Identifiers (PIDs); PID-Graph; Enhanced Publications; Scientific Information Services; Repositories (IR); Agile-software-development; Research Information Systems (CRIS).

ROR_ID <https://ror.org/008pnp284>ORCID_ID <https://orcid.org/0000-0002-5228-1970>**Balashova, Yuliya B.**

85

Yuliya B. Balashova, PhD, is a Professor at Saint Petersburg State University, Russia, and the Head of the Master's program Popular Science Journalism. She has taught at universities in Italy and Spain, and as a Fulbright scholar affiliated with Michigan State University, USA. She is the author of almost 100 scientific papers, including two monographs, on the history of literature and journalism, science communication, and education. Her current academic focus is interdisciplinary historical research. Her recent publications include *Debate about Scientific Popularization in the Russian Public Sphere (Based on Grey Literature Material)* (2017) and *Traditions of Russian Popular Science Journalism in the Socio-Cultural Context* (2018).

ROR_ID <https://ror.org/023znxa73>ORCID iD: <https://orcid.org/0000-0002-2315-8076>Email: j.balashova@spbu.ru**Baxter, David**

67, 117

David Baxter is the Information Specialist at Gambling Research Exchange (GREO). He is responsible for the collection development of GREO's Evidence Centre, a digital library of information resources, research datasets, and other grey literature that help advance knowledge of gambling-related issues. David sits on GreyNet's LIS Education and Training Committee, and is interested in the role of grey literature as a source of evidence for gambling research and policy. David holds a Master of Information from University of Toronto and a Bachelor of Science Honours in Biology from McMaster University.

ROR_ID <https://ror.org/047mprv74>ORCID_ID <https://orcid.org/0000-0001-5235-6728>E-mail david@greo.ca**Biagioni, Stefania**

45

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**Carlesi, Carlo**

45

Carlo Carlesi, graduated in Computer Science, worked since 1970 at the IEI (now ISTI) of the CNR in Pisa. He is currently a Research Associate of the Institute ISTI and he is involved in the following projects: PUMA - Publication Management. The Digital Library service allows public access (when permitted) through Internet to the published documents produced by CNR Organizations. And GreyGuide, portal and repository of good practice and resources in the field of grey literature.

ROR_ID <https://ror.org/05kacka20>ORCID iD <https://orcid.org/0000-0001-9808-6268>Email: carlo.carlesi@isti.cnr.it

Author Information *CONTINUED*

Farace, Dominic**45, 117**

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@greynet.org**Gelfand, Julia****55**

Julia Gelfand is the Applied Sciences, Engineering & Public Health Librarian at the University of California, Irvine Libraries where over the last 35 years she has performed many roles. She is active professionally and currently is a member of the Board of Directors of the Association of College & Research Libraries (ACRL), a division of the American Library Association, a member of the Science & Technology Section of the International Federation of Library Associations (IFLA) and is Secretary of Section T of the American Association for the Advancement of Science (AAAS). She writes and presents frequently on topics related to Scholarly Communication, Collection Management, Digital Scholarship, integration of multimedia in scientific literature, grey literature, social media, library as publisher. A graduate of Goucher College with graduate degrees from Case Western Reserve University, she is the recipient of many awards including the first GreyNet Award presented in 1999 and has been a Fulbright Fellow and a Thomas J. Watson Fellow. ROR_ID

<https://ror.org/04gyf1771>ORCID_ID <https://orcid.org/0000-0003-0424-497X>Email: jgelfand@uci.edu**Giannini, Silvia****117**

Silvia Giannini graduated and specialized in library sciences. Since 1987 she has been working in Pisa at the Institute for the Science and Technologies of Information "A. Faedo" of the Italian National Council of Research (ISTI-CNR) as a librarian. She is a member of the ISTI Networked Multimedia Information Systems Laboratory (NMIS). She is responsible of the library automation software "Libero" in use at the CNR

Research Area in Pisa and coordinates the bibliographic and managing activities of the ISTI library team. She cooperates in the design and development of the PUMA (Publication Management) & MetaPub, an infrastructure software for institutional and thematic Open Access repositories of published and grey literature produced by CNR.

Email: silvia.giannini@isti.cnr.itORCID_ID <https://orcid.org/0000-0001-7323-3786>ROR_ID <https://ror.org/05kacka20>**Kergosien, Eric****17**

Eric Kergosien, MCU since 2014, is working at the University Lille in the GERiiCO laboratory. His main area are Knowledge management and Text Mining applied to heterogeneous textual data with a special interest in semantic and geographic resources. One aspect of his work is Information Extraction with NLP and Text mining methods in several fields (scientific and technical data, libraries, land-use planning and medical). A second important aspect of his work is domain ontology engineering.

ROR_ID <https://ror.org/02kzqn938>ORCID_ID <https://orcid.org/0000-0002-2397-5519>Email: eric.kergosien@univ-lille.fr**Lipinski, Tomas A.****117**

Professor Lipinski completed his Juris Doctor (J.D.) from Marquette University Law School, Milwaukee, Wisconsin, received the Master of Laws (LL.M.) from The John Marshall Law School, Chicago, Illinois, and the Ph.D. from the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Dr. Lipinski has worked in a variety of legal settings including the private, public and non-profit sectors. He is the author of numerous articles and book chapters; his monographs include, THE LIBRARY'S LEGAL ANSWER BOOK co-authored with Mary Minow (2003); the COPYRIGHT LAW IN THE DISTANCE EDUCATION CLASSROOM (2005), THE COMPLETE COPYRIGHT LIABILITY HANDBOOK FOR LIBRARIANS AND EDUCATORS (2006), and THE LIBRARIAN'S LEGAL COMPANION FOR LICENSING INFORMATION RESOURCES AND SERVICES (2012). Recent articles and chapters include, Click Here to Cloud: End User Issues in Cloud Computing Terms of Service Agreements, in CHALLENGES OF INFORMATION MANAGEMENT BEYOND THE CLOUD: 4TH INTERNATIONAL SYMPOSIUM ON INFORMATION MANAGEMENT IN A CHANGING WORLD, IMCW 2013 (Revised Selected Papers.), with Kathrine Henderson, Hate Speech: Legal and Philosophical Aspects, in THE HANDBOOK OF INTELLECTUAL FREEDOM CONCEPTS (2014), in 2013 with Andrea Copeland, Look before you License: The Use of Public Sharing Websites in building Patron Initiated Public Library Repositories, PRESERVATION, DIGITAL TECHNOLOGY & CULTURE and in 2012, Law vs. Ethics, Conflict and Contrast in Laws Affecting the Role

Author Information *CONTINUED*

of Libraries, Schools and other Information Intermediaries, JOURNAL OF INFORMATION ETHICS. He has been a visiting professor in summers at the University of Pretoria-School of Information Technology (Pretoria, South Africa) and at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Professor Lipinski was the first named member of the Global Law Faculty, Faculty of Law, University of Leuven (Katholieke Universiteit Leuven), Belgium, in Fall of 2006 where he continues to lecture annually at its Centers for Intellectual Property Rights and Interdisciplinary Center for Law and ICT. He is active in copyright education and policy-making, chairing the ACRL Copyright Discussion Group, a member of the ALA OITP Committee on Legislation Copyright Subcommittee, a member of the Copyright and Other Legal Matters Committee of IFLA and serves as an IFLA delegate to the World Intellectual Property Organization's Standing Committee on Copyright and Other Rights. In October of 2014 he returned to the University of Wisconsin—Milwaukee to serve as Professor and Dean of its i-School, the School of Information Studies.

ROR_ID <https://ror.org/031q21x57>

Email tlipinsk@uwm.edu

Marshall, Andrea 89

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: andrearieithmarshall@gmail.com

Memari Hanji, Hamideh 93

Hamideh Memari Hanji is Senior Administrator of ISSN NC Iran and member of ISSN International Centre GB, National Library and Archives of Iran, Tehran, Iran

Email: hmeamari@yahoo.com

Molino, Anna 117

Anna Molino graduated in Linguistics at the University of Pisa in 2010. Since 2013, she works at ISTI-CNR ("Istituto di Scienza e Tecnologie dell'Informazione - A. Faedo"). She worked as project assistant and financial manager in various EU funded and national research projects for the Digital Libraries group of the NeMIS lab. Since 2018, she joined the Library staff of the CNR Area in Pisa. She is actively involved in all library duties as well as in the practice of Open Access and Open Science principles. She also contributes in the language revision and translation of the scientific production of the library. ROR_ID <https://ror.org/05kacka20>

ORCID_ID <https://orcid.org/0000-0002-4793-3727>

Email: anna.molino@isti.cnr.it

Palcullo, Vince Ervin V. 101

Vince Ervin V. Palcullo, RL, is the current Archives/Special Collection Librarian and a part-time LIS faculty of Central Philippine University, Iloilo City, Philippines. His research interests are in Information Behaviour, Disaster Management, Citation Analysis, Systematic Review, and History and Culture.

ORCID_ID <https://orcid.org/0000-0001-9680-3702>

Email: vevpalcullo@cpu.edu.ph

Peleña, Ma. Cynthia T. 101

Ma. Cynthia T. Peleña is currently the Director of Libraries, Central Philippine University. Before her appointment, she served as program coordinator of both Master in Library Science with Specialization in Theological Librarianship (MLIS-TL), Master in Library and Library Science (MLIS) and a Bachelor of Library Science (BLIS) from 2013-2019 and a part-time faculty member under the School of Graduate Studies, College of Education, and the College of Computer Studies Central Philippine University. She graduated with a Master in Library and Information Science from Central Philippine University in 2010. In 2017, Cynthia received a professional enhancement grant as a visiting professor at Columbia Theological Seminary, Decatur, Georgia, USA from July-September, 2017. Currently, she is one of the task force members of the International Theological Librarianship Education Task Force (ITLE) under the American Theological Library Association from 2018 to the present.

ORCID_ID <https://orcid.org/0000-0002-1121-5029>

Plank, Margret 33

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

Author Information *CONTINUED*

Potočnik, Veronika

117

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 <https://ror.org/04tjx1w16>

Email: veronika.potocnik@nuk.uni-lj.si

Prost, Hélène

17

Hélène Prost is information professional at the Institute of Scientific and Technical Information (CNRS) and associate member of the GERiICO research laboratory (University of Lille 3). She is interested in empirical library and information sciences and statistical data analysis. She participates in research projects on evaluation of collections, document delivery, usage analysis, grey literature and open access, and she is author of several publications.

ROR_ID <https://ror.org/02mn0vt57>

ORCID_ID <https://orcid.org/0000-0002-7982-2765>

Email: helene.prost@inist.fr

Sadat Hashemi, Somayeh

111

Somayeh Hashemi is currently in charge of Information Services Department of the National Library of Iran. As senior reference librarian she has gathered valuable experience in several large-scale projects. She has implemented telegram messenger as a new virtual reference services. She has been working in the National Library and Archives of Iran for the past 18 years in different positions.

Savić, Dobrica

11, 117

Dr Dobrica Savić is an information and knowledge management consultant at the IAEA. He holds a Doctorate degree from Middlesex University in London, an MPhil degree in Library and Information Science from Loughborough University, UK, an MA in International Relations from the University of Belgrade, Serbia, as well as a Graduate Diploma in Public Administration, Concordia University, Montreal, Canada. He has extensive experience in the management and operations of web, library, information and knowledge management, as well as records management and archives services across various United Nations agencies, including UNV, UNESCO, World Bank, ICAO, and the IAEA. His main interests are digital transformation, creativity, innovation, and the use of IT for successful information and documentation services.

ROR_ID <https://ror.org/00gtfax65>

ORCID_ID <https://orcid.org/0000-0003-1123-9693>

Contact: www.linkedin.com/in/dobricasavic

Schöpfel, Joachim

17, 117

Joachim Schöpfel 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 NDLTD.

ROR_ID <https://ror.org/02kzqn938>

ORCID_ID <https://orcid.org/0000-0002-4000-807X>

Email: joachim.schopfel@univ-lille.fr

Thiault, Florence

17

Senior lecturer in Library and Information Science at University Rennes 2, laboratory PREFICS. Co-manager of URFIST (regional unit of training in scientific and technical information) of Rennes.

ROR_ID <https://ror.org/01m84wm78>

ORCID_ID <https://orcid.org/0000-0001-7466-5855>

Email florence.thiault@univ-rennes2.fr



Twenty-Fourth International Conference on Grey Literature Publishing Grey Literature in the Digital Century

National Library of Medicine
Lister Hill Auditorium, 5-6 December 2022
Bethesda, Maryland USA

Conference Announcement



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 will address the components of such shared workflows embedded in FAIR data principles and implemented by diverse communities of practice in this our digital Century.

Related Conference Topics

e-Publishing – *AV-Portals, Blogs, Data Archives, Repositories, etc.* **Digital Publications** – *Clinical Trials, ETDs, Data Papers, Preprints, etc.* **Research and Metadata** – *Datasets, Databases, PIDS, DOI, ORCID, ROR-ID, etc.* **Stakeholders and Policies** – *Content and Service Providers, Authors and Researchers, OA, CC, FAIR, etc.* **Communities of Practice** – *Agriculture, Biomedicine, Economics, Fisheries, Physics, Library and Information, etc.*

GL2022 Conference Dateline

Mar. 31	April 11	April 20	April 25	Nov. 1 st	Nov.15	Dec. 5-6
Call for Papers Closes	Program Committee Meeting	Author and Session Notifications	Call for Posters Opens	Early Registration Closes	Conference Papers and Posters Due	GL2022 Conference Convenes

On behalf of *GreyNet* and the International Conference Series on Grey Literature, 1992–2022

Index to Authors

A-B

Aminalroaya, Roya	111
Baars, Chris	45
Balashova, Yuliya B.	85
Baxter, David	67, 117
Biagioni, Stefania	45

C-D- F

Carlesi, Carlo	45
Dumenden, Via Marie F.	101
Farace, Dominic	45, 117

G-H

Gelfand, Julia	55
Geromiano, Joy F.	101
Giannini, Silvia	117
Hakhamaneshi, Nooshin	111

K-L

Kergosien, Eric	17
Kösters, Jens	33
Lipinski, Tomas A.	117

M

Marshall, Andrea	89
Memari Hanji, Hamideh	93
Molino, Anna	117

N-P

Nour, Nika	55
Palcullo, Vince Ervin V.	101
Peleña, Ma. Cynthia T.	101
Plank, Margret	33
Potočnik, Veronika	117
Prost, H��l��ne	17

R-S-T

Rezvani, Ehsan	111
Rivera, Donna May C.	101
Sadat Hashemi, Somayeh	111
Savi��, Dobrica	11, 117
Sch��pfel, Joachim	17, 117
Daryl Superio	101
Thiault, Florence	17

Managing Grey Literature

Technical Services Perspectives

Edited by Michelle Leonard & Susan E. Thomas

Using this guide, collection managers and acquisitions librarians, preservation librarians, cataloguers, and library managers will understand how to utilise the technical services workflow to process and showcase this unique material.

An important resource for scholarly research, grey literature is relevant to every discipline. It's also often more current than commercial publications. Unfortunately, though it provides a richness of content, this type of scholarly resource is often overlooked when conducting research. This book aims to change that, describing the importance of grey literature and offering a holistic approach to successfully integrating it into library collections.

Readers will learn:

- An overview of grey literature that discusses its importance to researchers, scholars, and students.
- Collections policies for selection and deselection, complete with a suggested workflow.
- Information about vendors, OA, and other aspects of acquisitions.
- Methods for promoting grey literature in library collections, including institutional repositories.
- Tips for marketing, branding, outreach, and best communication practices for colleagues, administrators, and patrons.

Table of Contents:

Chapter 1	Grey Literature: its Past, Present and Future – Leonid Pavlov
Chapter 2	Managing and Weeding the Grey Literature Collection – Hillary Fox and Cynthia Levine
Chapter 3	Ideas and Challenges in Cataloging Grey Literature – Rachel Berman Turner
Chapter 4	Persistent Identifiers and Grey Literature: A PID Project and GreyNet Use Case – Dominic Farace, Stefania Biagioni, Carlo Carlesi, Chris Baars
Chapter 5	Communicating the Value of International Grey Literature – Tamsin Vicary, Maria Kalentsits, Florine Lim, Daryl Superio
Chapter 6	Impact of Digital Transformation on Grey Literature – Dobrica Savić



9780838948811, Dec 2021

Paperback, 218pp

£71.50 / €81.00

ALCTS Monograph

ALA Editions

American Library Association

ALA American
Library
Association

FURTHER INFORMATION

Tel: +44(0)20 7240 0856
Fax: +44 (0)20 7379 0609
Email: info@eurospan.co.uk

CUSTOMER SERVICES

Tel: +44 (0)1767 604972
Fax: +44 (0)1767 601640
Email: eurospan@turpin-distribution.com

Eurospan Ltd
Gray's Inn House
127 Clerkenwell Road
London, EC1R 5DB, UK