e-ISSN 1574-180X

An International Journal on Grey Literature

Volume 19, Number 3, Autumn 2023

'CURRENT DEVELOPMENTS AND INITIATIVES INVOLVING GREY LITERATURE'

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

The Grey Journal

An International Journal on Grey Literature

COLOPHON

About TGJ

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

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

CIP
The Grey Journal (TGJ) : An international journal on grey literature / Dominic Farace (Journal Editor); Jerry Frantzen (Technical Editor) ; GreyNet International, Grey Literature Network Service. - Amsterdam: TextRelease, Volume 19, Number 3, Autumn 2023 – TIB (DE), CVTISR (SK), EBSCO (USA), KISTI (KR), NIS-IAEA (UN), NTK (CZ), Figshare (UK) and the University of Florida (USA) are Corporate Authors and Associate Members of GreyNet International.
e-ISSN 1574-180X (PDF)

Subscription Rate:
€240 institutional

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

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

About GreyNet

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

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

Contents

'Current Developments and Initiatives Involving Grey Literature'

PsycEXTRA®

EXPAND BEHAVIORAL SCIENCE RESEARCH WITH THE PREMIER RESOURCE FOR GRAY LITERATURE

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




Explore original, cutting-edge, high-quality research

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

Available via EBSCO

EDITOR'S NOTE

By drawing public awareness to grey literature without at the same time incorporating the results of studies over the past thirty years would be counterproductive. A campaign of public awareness with a new focus on grey literature is required. However, such an initiative can only be achieved in collaboration with the many grey literature stakeholders led by the publishers in grey literature, who are on the forefront of research and innovation in their specific domains of science.

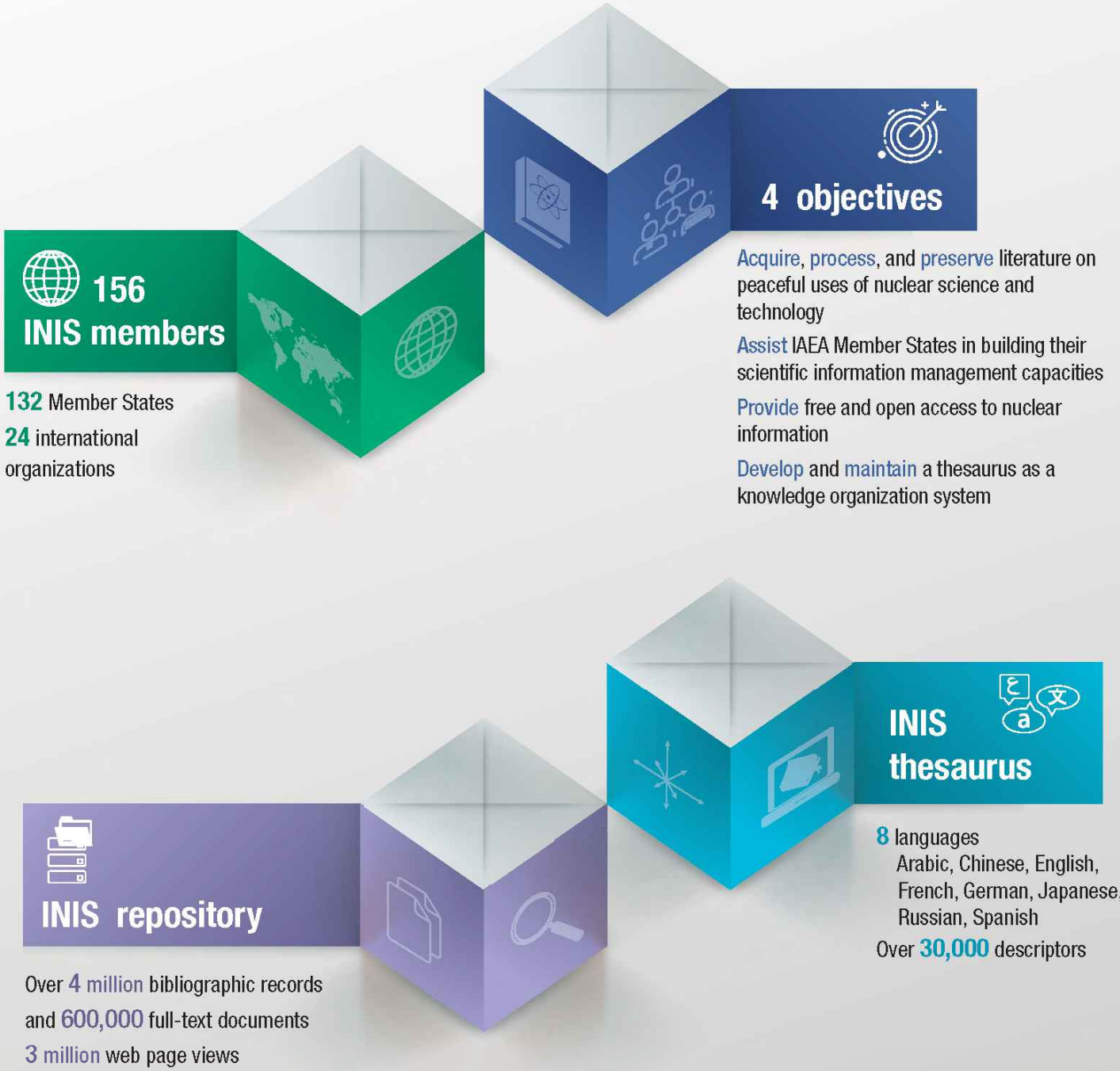
GreyNet International owing to the support of its associate members has over the past three decades sought to promote and undertake research, publication, open access, education, and public awareness of grey literature. Emphasis on its supply side was always its starting point from which the field could confront the problems voiced by the demand side and adequately address them.

Speaking on behalf of GreyNet's community of practice, I maintain that by formalizing a consortium of grey literature publishers, barriers can be overcome. An initiative undertaken earlier this year with the PUBGREY Registry can be seen as a first step in this direction. The time appears right and the opportunity presents itself during the upcoming 25th International Conference on grey literature this November in Amsterdam.

Dominic Farace
Journal Editor

International Nuclear Information System – INIS

Organizing the world's nuclear information and making it universally accessible



156 INIS members

132 Member States

24 international organizations

4 objectives

Acquire, **process**, and **preserve** literature on peaceful uses of nuclear science and technology

Assist IAEA Member States in building their scientific information management capacities

Provide free and open access to nuclear information

Develop and **maintain** a thesaurus as a knowledge organization system

INIS thesaurus

8 languages
Arabic, Chinese, English, French, German, Japanese, Russian, Spanish

Over **30,000** descriptors

INIS repository

Over **4 million** bibliographic records and **600,000** full-text documents

3 million web page views

2 million unique searches

1 million unique visitors

IAEA | International Nuclear Information System INIS

www.iaea.org/inis

February 2019
19-00898

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

Brian Bales,
Nuclear Information Section, International Atomic Energy Agency

Abstract

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

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

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

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

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

Current Situation

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

A well-developed repository may be defined as one that best meets the interests of its sponsoring organization. By extension and considering that a sponsoring organization is likely to want its users to also find the product useful, the interests and needs of end users should also be considered.

[*] First published in the GL2022 Conference Proceedings, February 2023
Video Presentation: https://av.tib.eu/media/59873

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

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

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

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

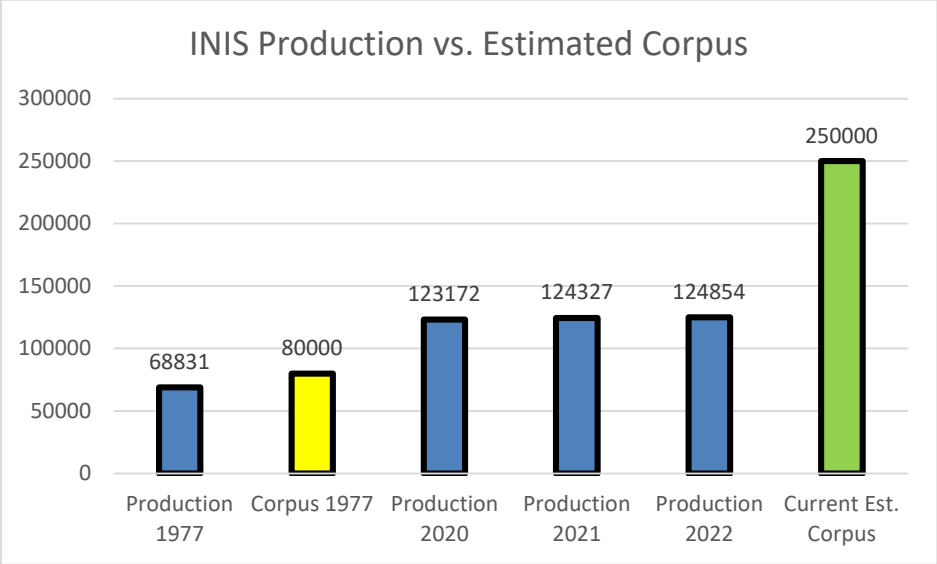


Figure 1: INIS 1977 Production vs. Estimated 1977 Corpus, Current Production vs Estimated Current Corpus

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

Exemplars

The Astrophysics Data System[5]

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

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

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

INSPIRE-HEP[7]

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

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

Standards

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

The FAIRsFAIR project "Fostering FAIR Data Practices in Europe"[9], seeks to establish its principles within what it calls the "infrastructure of science." These principles are that scientific information and data should be findable, accessible, interoperable, and re-usable. Much of this speaks to the need for openness. Openness removes financial barriers to science, where many publishers had been and continue to charge high access fees.[10]

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

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

Characteristics

From studying the needs of INIS, as well as the configurations of exemplar repositories and

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

Timeliness

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

Openness

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

Preservation

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

User Friendliness

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

Comprehensiveness

In essence, each of these is about extending availability. Timeliness and preservation extend the availability of records temporally. Openness extends availability physically and financially, by

making records available in multiple locations and without cost. User friendliness makes records available to both novice and expert users with a compelling interface. Comprehensiveness extends the availability of records to all possible candidate records within the scope of the repository.

Future Work

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

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

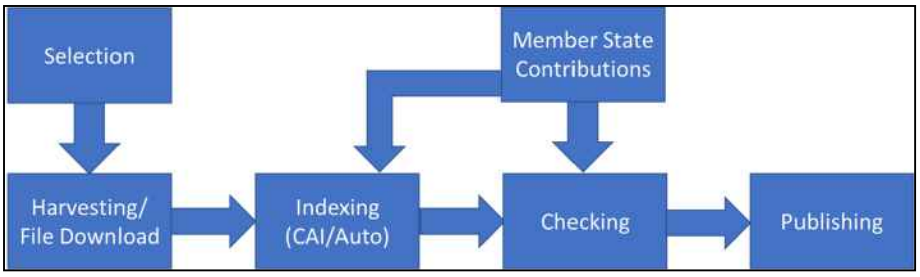


Figure 2: Simplified Current Workflow, Most Steps Require Manual Initiation, Capable of 125 000 to 130 000 per year

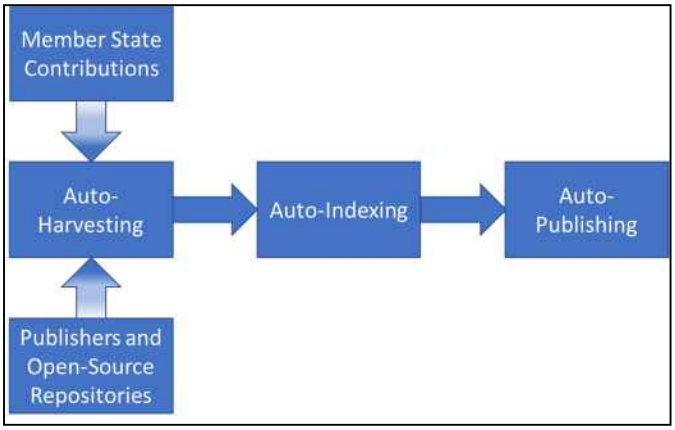


Figure 3: Proposed Future Workflow, All Steps Automated

Conclusion

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

References

[1] Directory of Open Access Repositories, "Statistics", https://v2.sherpa.ac.uk/view/repository_visualisations/1.html

[2] International Nuclear Information System (INIS), "Home", https://www.iaea.org/resources/databases/inis

[3] INIS Repository Search, https://inis.iaea.org/search/

[4] Ivan S. Zheludev and Hans W. Groenewegen, "INIS: The International Nuclear Information System", IAEA Bulletin, Vol. 20, No. 4, https://www.iaea.org/sites/default/files/publications/magazines/bulletin/bull20-4/20405060717.pdf

[5] The Astrophysics Data System, https://ui.adsabs.harvard.edu/

[6] The Astrophysics Data System, "About", https://ui.adsabs.harvard.edu/about/

[7] INSPIRE-HEP, https://inspirehep.net/

[8] ISNPIRE-HEP, "About", https://help.inspirehep.net/knowledge-base/about-inspire/

[9] FAIRSFAIR, https://www.fairsfair.eu/

[10] European Commission, "Turning FAIR into reality", https://op.europa.eu/s/xoR2

[11] CoreTrustSeal, https://www.coretrustseal.org/

[12] CoreTrustSeal, "Trustworthy Data Repositories Requirements", https://www.coretrustseal.org/why-certification/requirements/

[13] FAIRSFAIR, "CoreTrustSeal + FAIRenabling Capability Maturity Model", https://www.fairsfair.eu/node/797/pdf

[14] OpenAIRE, https://www.openaire.eu/

[15] OpenAIRE, "Guidelines", https://guidelines.openaire.eu/en/latest/

A Retrospective on the Challenges of Incorporating Grey Literature into a Scholarly Publishing Platform*

Alistair Reece, GeoScienceWorld, United States

Abstract

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

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

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

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

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

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

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

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

Introduction

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

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

* First published in the GL2022 Conference Proceedings, February 2023
Video Presentation: https://av.tib.eu/media/59862

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

Background

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

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

Content Preparation

- Implications for Search
- New Business Models

In the retrospective we asked the following questions:

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

Content Preparation

What went well?

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

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

Having a reliable content manifest that included content types we knew that there would be development work necessary to support the loading and ingestion of the Grey Literature elements of the content set. Identifying the scope and types of Grey Literature within the set allowed us to avoid many of the unexpected surprises of a content migration. As such, GeoScienceWorld has implemented a thorough content discovery process into our acquisition workflow so that we have a proper understanding of the content being brought into our aggregation. This process is being applied to all new content acquisitions regardless of the content type involved, and has allowed us to better plan project timelines and set expectations with new publishers.

What could be better?

There was one area in particular with regards to the content preparation process that we really could have done much better, and that was having an awareness of the absence of a common Document Type Definition (DTD) for creating XML files to load Grey Literature. Within traditional scholarly publishing there exist standard DTDs for both journal articles and ebooks, JATS and BITS respectively. These DTDs are used by all the major publishing platforms for loading content,

making a certain level of interoperability available during migrations from one platform to another. We were unaware that such a shared standard does not exist for Grey Literature.

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

What have we learned?

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

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

What questions remain?

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

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

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

Implications for Search

What went well?

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

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

The current GeoScienceWorld search experience includes such visual clues when it comes to the authorization state of a piece of content, showing the user if they have the authorization to view with a green check mark or a shopping cart icon if not. The search experience also uses visual cues to identify Open Access content through the use of an unlocked orange padlock. Knowing that such visual aids are useful to the user, GeoScienceWorld knows that a similar set of cues will be required to identify the peer-review status of the content. We are aware of sets of icons

already available for this kind of cue and will implement them in our search results page when we have Grey Literature as part of the platform.

What could be better?

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

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

What have we learned?

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

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

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

As GeoScienceWorld has spent a lot of time as a result of the Grey Literature project investigating how search functions and is presented on our platform, we have realized that search results contain overwhelming amounts of metadata. Each search result presents the user with at least eight metadata items, as well as a snippet of the content, where the search term is found. The search result also includes options to view the abstract, download the PDF, purchase the content, or add to the Citation Manager tool. Based on feedback from users, we have realized that half the metadata fields displayed are not relevant to the search itself, and instead serve to clutter the results and overwhelm the user.

What questions remain?

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

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

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

New Business Models

What went well?

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

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

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

What could be better?

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

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

Despite identifying our primary market, and having a method for potentially uncovering unexpecting secondary markets, GeoScienceWorld is keen to find a unique selling point for our academic customers, who form the majority of our customer base. We are convinced that this Grey Literature set has value to the researchers working in the academic institutions that we serve.

What have we learned?

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

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

We also learned that our end users, the research community as opposed to the librarian community, are less concerned about the peer-reviewed status of the content than they are about the value and reproducibility of the science behind the article itself. We suspect that this

may be a behavior specific to researchers within the geosciences, and are seeking to validate this assumption.

What questions remain?

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

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

Conclusions

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

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

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

We are convinced that Grey Literature has very real, tangible, value to the research community that uses the traditional peer-reviewed literature that forms the basis of our platform. That value makes Grey Literature repositories an attractive proposition in our content acquisition efforts.

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

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

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

Perhaps our biggest conclusion as a result of this project so far, and the retrospective we undertook, is that there is a distinct need for leadership within the Grey Literature community when it comes to metadata preparation for content loading and presentation standards. As more and more traditionally scholarly publishers bring non-journal and book chapter content

into the platforms that we already use, there is an opportunity for the Grey LIterature community to work together with the platforms to define shared standards to make Grey Literature a more accepted part of the research databases that science relies upon.

References

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

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

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

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

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

Introduction

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

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

Method of Approach

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

Survey Questions

1. Do you consent to take part in this study on the terms described above in the Participant Consent Form?
2. What is your organization's name and (if applicable) the acronym?
3. What is the URL of your organization's homepage?
4. Has your organization published one or more documents on the 2022 War in Ukraine?
5. If so, please provide a hyperlink or DOI to one of the publications?
6. What document type best describes the publication? (e.g., Blog, Podcast, Report, Policy Document, etc.)
7. Does the publication contain research data?
8. Does the publication in whole or part consist of audio-visual material?
9. Are you an author, co-author, or editor of the publication?
10. Please enter your name and/or ORCiD along with your email address?

* First published in the GL2022 Conference Proceedings, February 2023
Video Presentation: https://av.tib.eu/media/59876

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

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

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

Part One

Overview of the Survey Findings

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

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

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

Summary of the Respondents' Comments

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

When asked if their organization published one or more documents on the War in Ukraine 2022, those who answered 'yes', further commented that their library provided information services in Ukrainian as well as basic information about help for Ukrainian refugees in their country. Research grants were made available to Ukrainian researchers. Reports of statements by academic and religious spokesmen were published. Reliable curated links regarding the War in Ukraine were shared. Events were organized on campuses dealing with the War. Appeals to end the War were drawn-up, signed, and published by both researchers and staff. Other respondents mentioned that while they did not publish, they had aggregated a manifold number of reports on the War in Ukraine.

Those who answered 'no' to the question responded that while their department had not yet published about the War in Ukraine, perhaps others on their campus did. Mention was made that technically they were not publishers. Further mention was made that their repository did

however contain published articles dealing with that which led up to the War in Ukraine. Other respondents indicated that they intended to publish on the War in Ukraine in the near future.

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

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

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

Summary of the Primary Texts

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

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

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

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

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

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

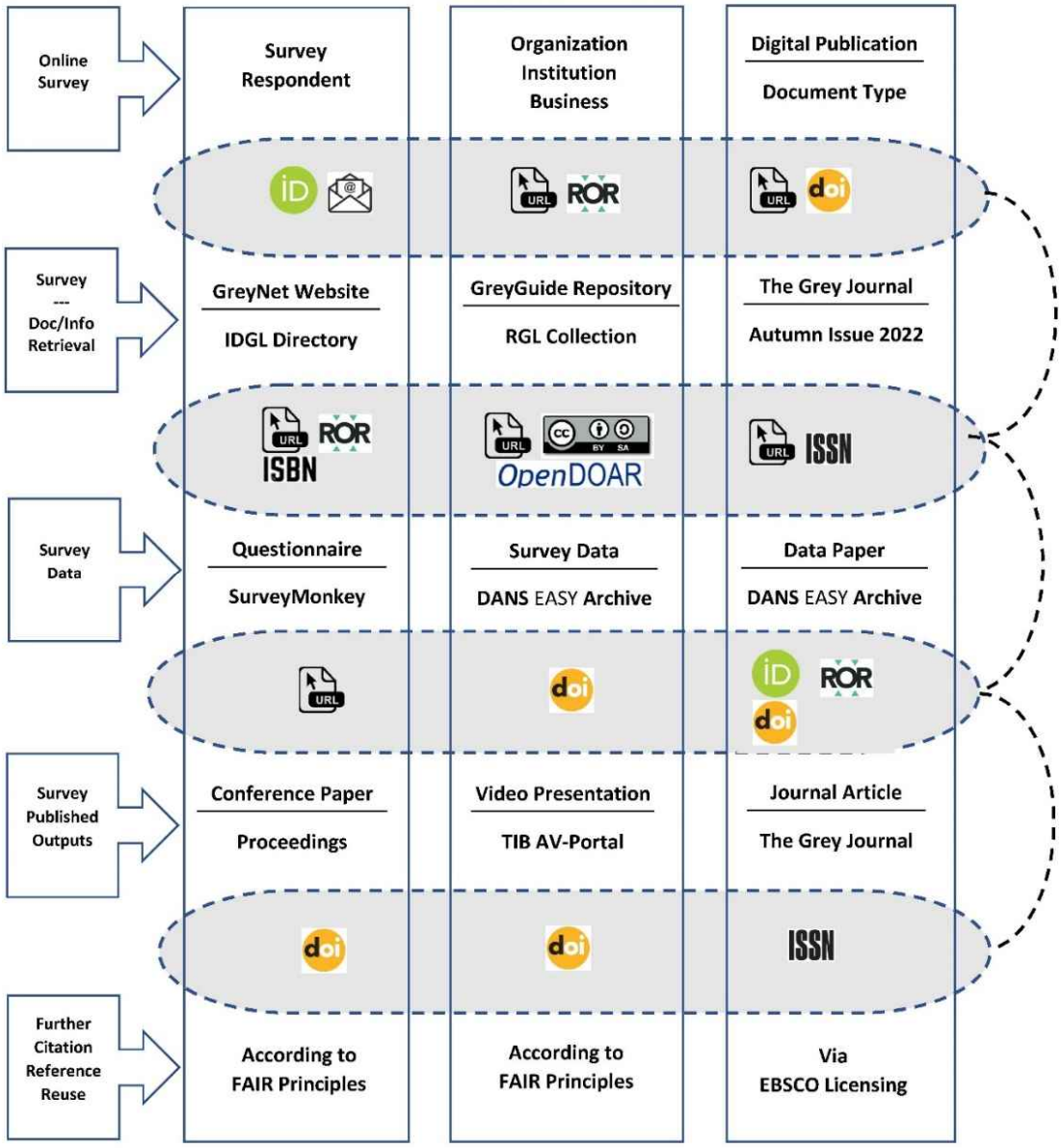
Early on in the War, publications discussed the devastating effect it would have on the cultural heritage of Ukraine. Already a museum housing the work of an UNESCO recognized artist was destroyed. SUCHO, an initiative of over 1,500 international volunteers that collaborate online to

digitize and preserve Ukrainian cultural heritage published a digital poster on their efforts. Likewise, early on in the conflict, FAO signals the risks involved for global agricultural markets that would aggravate the world food crisis.

Part Two

Tracking Digital and Product Identifiers in Publishing

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



The schematic diagram (*above*) shows GreyNet's publications both in textual and non-textual formats originating with the online survey and connected throughout by actionable persistent identifiers and other product identifiers. Beginning with the publication of the link to the online survey (*top left column*), the responses provide access to data related to the respondents, their organizations, the types of grey literature documents published on the War in Ukraine, etc. In turn, these data are themselves connected to digital and product identifiers such as ORCiDs, ROR IDs, URLs, and DOIs.

Impact of Survey Data on Digital Publishing

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

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

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

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

Concluding Observations

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

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

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

Linked References

[1] https://greynet.org/internationaldirectory.html

[2] https://doi.org/10.26069/greynet-2022-000.482-gg

[3] https://www.greynet.org/thegreyjournal/currentissue.html

[4] https://ror.org/01pxfxj80

[5] https://www.go-fair.org/fair-principles/

[6] https://surveymonkey.com/

[7] https://doi.org/10.17026/dans-zhs-fewk

[8] https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:254710/tab/2

[9] https://www.ebsco.com/products/research-databases/library-information-science-technology-abstracts-full-text

[10] http://greyguiderep.isti.cnr.it/listtitoligl.php?authority=GreyGuide&collection=GLP&langver=en&RighePag=100

[11] http://greyguide.isti.cnr.it/index.php/greyguideportal/document-share/gl-proceedings-1993

[12] https://av.tib.eu/publisher/GreyNet%20International

[13] https://www.sciencedirect.com/science/article/pii/S2666389920302440

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

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

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

1. Introduction

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

2. Study Design

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

* First published in the GL2022 Conference Proceedings, February 2023
Video Presentation: https://av.tib.eu/media/59861

Technical Information (SSSTI), the State and University Library of Gottingen (SUB), the U.S. Department of Energy Office of Scientific and Technical Information (OSTI), the U.S. National Library of Medicine, and VTT Technical Research Centre of Finland. These organizations were given a duration of two months (08/22/22-10/19/22) to complete the survey.

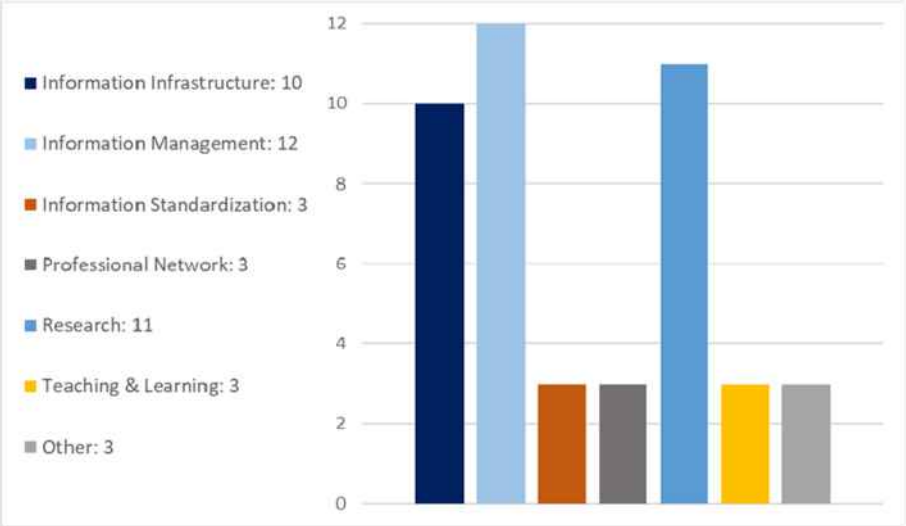
The total number of participant organizations that completed the survey was fifteen. Due to the small sample size, the results of this survey do not claim to be representative.

3. Results

3.1 Formal Criteria

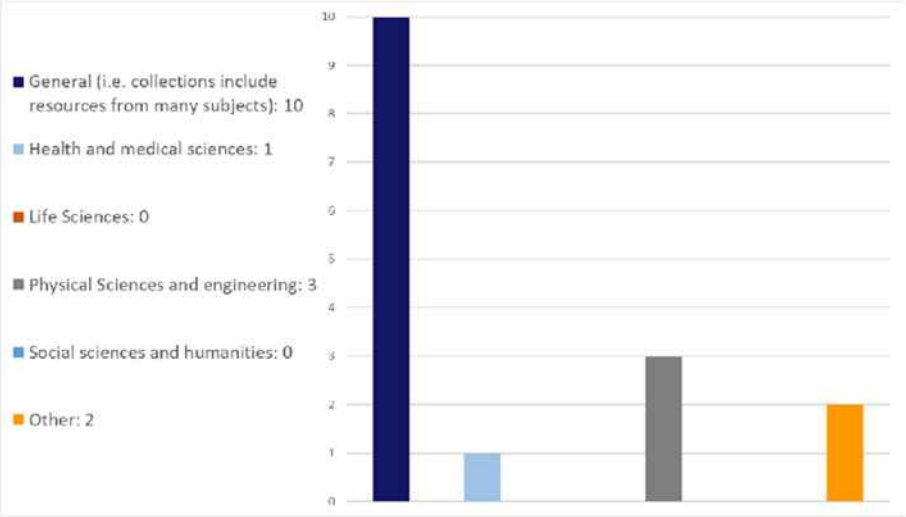
3.1.1 What are your organization's major functions or activities?

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



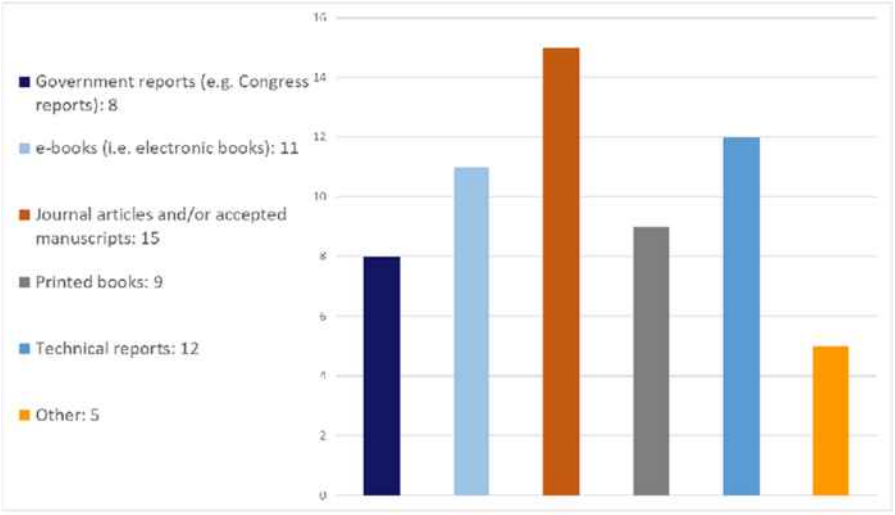
3.1.2 What is the primary subject focus of your organization?

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



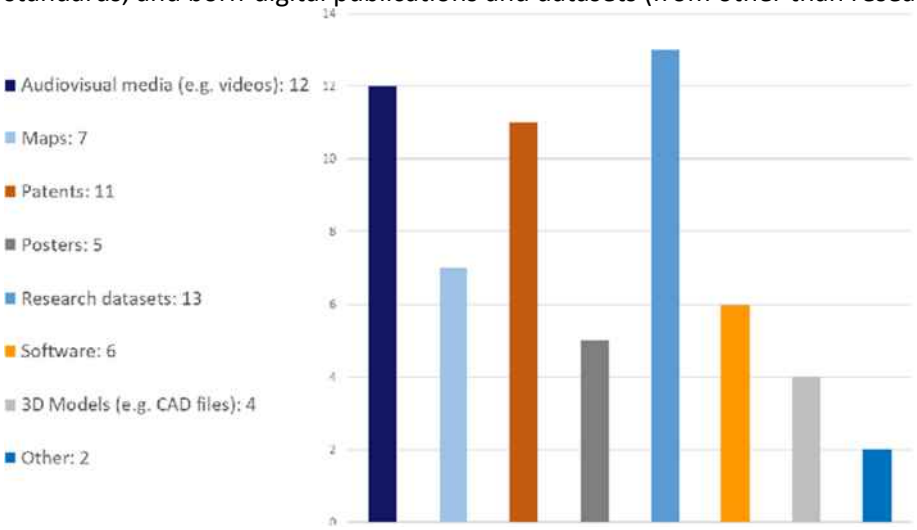
3.1.3 What types of digital textual scientific and technical information materials does your organization collect or store?

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



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

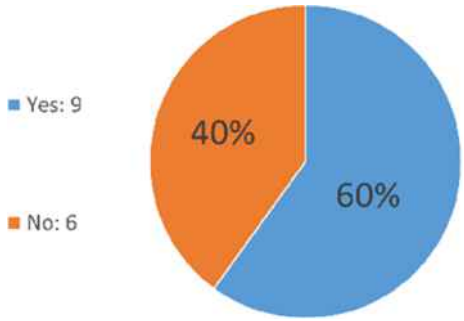
A wide range of object types were named. In the order of most frequently mentioned: research datasets (such as measurement data, material samples, structural formulas, simulation data, and geospatial data), audiovisual material, patents, maps, software, posters, 3D models (CAD) and others such as oral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



3.2 Audiovisual Media

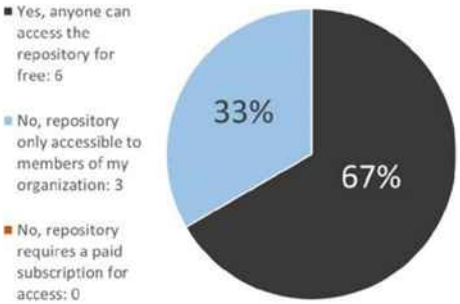
3.2.1 Does your organization host a repository for audiovisual media?

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



3.2.2 Is the audiovisual media repository publicly accessible?

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

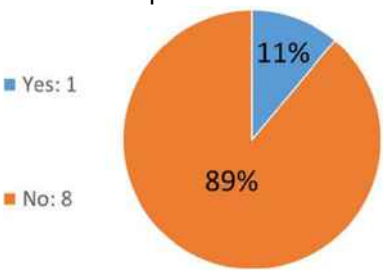


3.2.3 If your organization does not host a repository for audiovisual media, what, if any, external services are used to provide access?

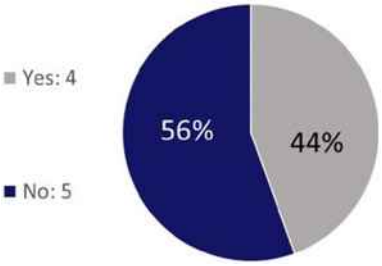
The named external services used to provide access to audiovisual media were Youtube, TikTok, Youku, WorldWideScience.org, J-Stage Data and NLM Digital Collections. Additionally, a collection of DVDs that can be accessed locally was mentioned.

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

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

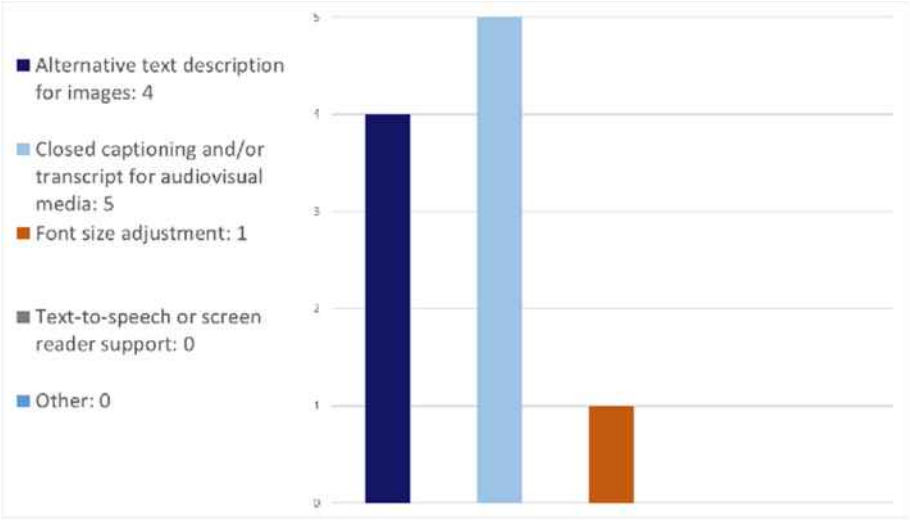


3.2.5 Does your organization provide long-term digital preservation for audiovisual media?
Four respondents answered that their organization provides long-term digital preservation, with five answering that they do not.

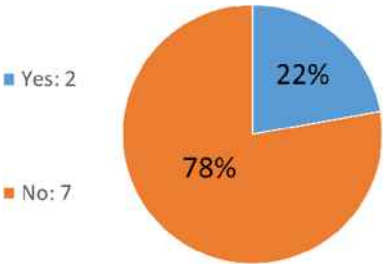


3.2.6 Does your organization provide any of the following infrastructures, tools, or services to make the audiovisual media repository accessible to users with disabilities?
Regarding the provision of tools and infrastructures for users with disabilities, the organizations provide services such as closed captioning or transcripts, alternative text description and text size adjustments for audiovisual material.

3.2.7 Does your organization provide any of the following training resources in relation to the management or publication of audiovisual media?
Most of the organizations provide written text or guidance, as well as instructional videos used as training resources in relation to management or publication of audiovisual media.



3.2.8 Does your organization assign persistent identifiers to audiovisual media?
Only two organizations answered that they assign persistent Identifiers for audiovisual media, those being specifically Digital Object Identifiers (DOIs).

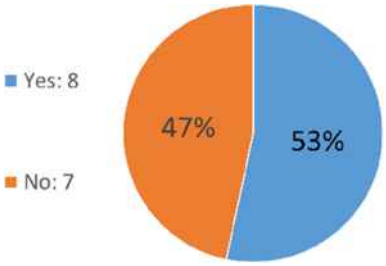


3.2.9 Which DOI registration services are used to assign DOIs to audiovisual media?
In addition to DataCite's DOI registration service, one organization utilizes the Entertainment Identifier Registry (EIDR).

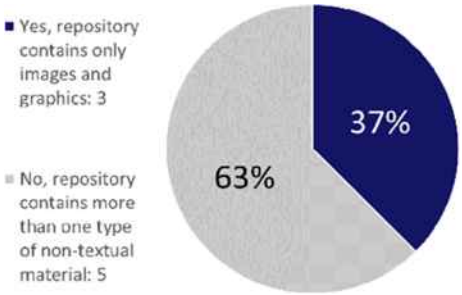
3.2.10 Does your organization assign any of the following types of persistent identifiers to audiovisual media (in addition to or instead of DOIs)?
None of the organizations use other persistent identifiers such as Arxiv or URN.

3.3 Images & Graphics

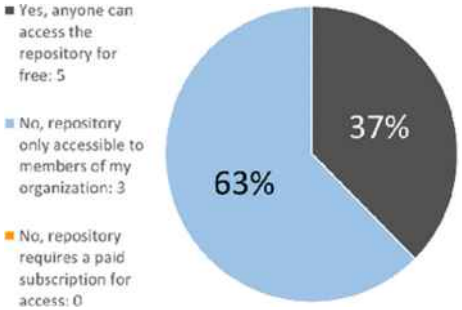
3.3.1 Does your organization host a repository for Images and Graphics?
Eight of the respondents answered that their organization had a repository for images and graphics, while the other seven stated that their organization did not.



3.3.2 Is the repository exclusive to Images and Graphics?
Five respondents claimed that their repository was not exclusive to images and graphics, and three replied that their repository was exclusive for images and graphics.



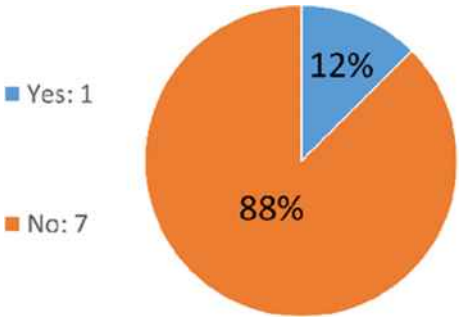
3.3.3 Is the repository for Images and Graphics publicly accessible?
Five organizations stated that their repository was publicly available, while three answered that the repository was only accessible to members of their organization.



3.3.4 If your organization does not host a repository for images and graphics, what, if any, external services are used to provide access?
The named external services used to provide access to images and graphics were Youtube, TikTok, Youku, J-Staged, Figshare, Zenodo and Wikicommons. Additionally, a collection of analogue maps that can be accessed locally was listed.

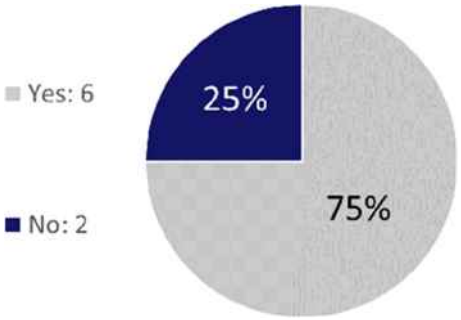
3.3.5 Does your organization provide a visual search tool for the images and graphics repository?

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



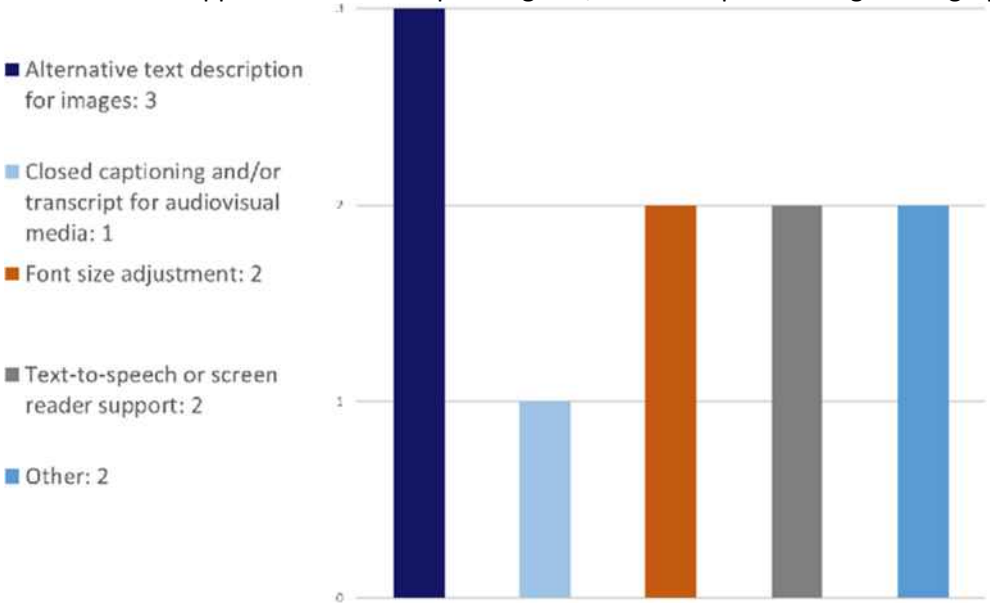
3.3.6 Does your organization provide long-term digital preservation for images and graphics?

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



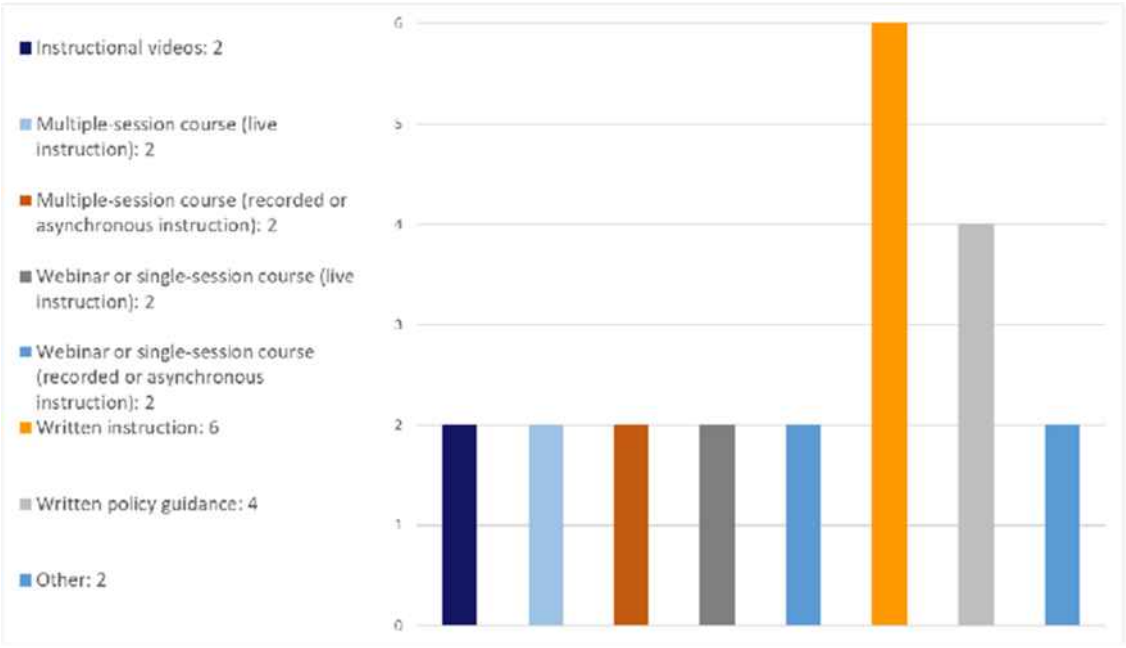
3.3.7 Does your organization provide any of the following infrastructures, tools, or services to make the images and graphics repository accessible to users with disabilities?

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



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

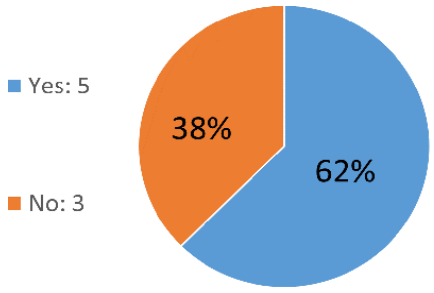
With respect to training materials, the most prevalent resource provided is written instruction and policy guidance.



3.3.9 Does your organization assign persistent identifiers to images and graphics?
Five organizations assign persistent identifiers to images and graphics.

3.3.10 Does your organization assign digital object identifiers (DOIs) to images and graphics? Which DOI registration services are used to assign DOIs to images and graphics?
Three organizations assign DOIs to images and graphics using either the DataCite or KISTI registration services.

3.3.11 Does your organization assign any of the following types of persistent identifiers to images and graphics (in addition to or instead of DOIs)?
The responses included internal persistent identifiers assigned to images and graphics.

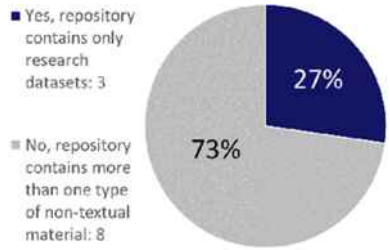
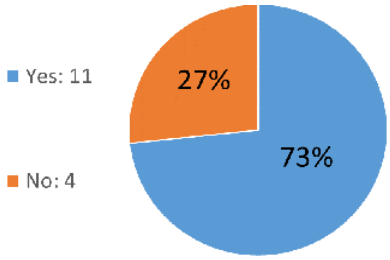
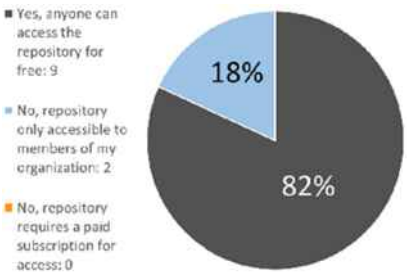
Does your organization assign persistent identifiers to images and graphics?	Does your organization assign digital object identifiers to images and graphics? • Yes: 3 • No: 2 Which DOI registration services are used to assign DOIs to images and graphics? • DataCite: 3 • Entertainment Identifier Registry (EIDR): 0 • ISTIC & Wanfang Data: 0 • Japan Link Center (JaLC): 0 • Korea Institute of Science and Technology (KISTI): 1 • Multilingual European DOI Registration Agency: 0 • Publications Office of the European Union: 0 • Other: 0 Does your organization assign any of the following types of persistent identifiers to images and graphics (in addition to or instead of DOIs)? • ARK (archival resource key: 0 • ArXiv: 0 • Internal identifier (i.e. organization specific identifier): 1 • URN (uniform resource name): 0 • Other: 1
	

3.4 Research Datasets

3.4.1 Does your organization host a repository for research datasets?
Eleven organizations answered that they do host a repository for research datasets, while four do not.

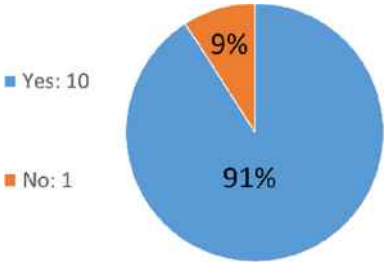
3.4.2 Is the repository exclusive to research datasets?
Eight out of the eleven organizations stated that their repository was not exclusive to research datasets; three of the organizations' repositories only contain research datasets.

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

<table>
<tr><td>Does your organization host a repository for research datasets?</td><td>Is the repository exclusive to research datasets?</td><td> 
</td></tr>
<tr><td> 
</td><td>Is the research datasets repository publicly accessible?</td><td> 
</td></tr>
</table>

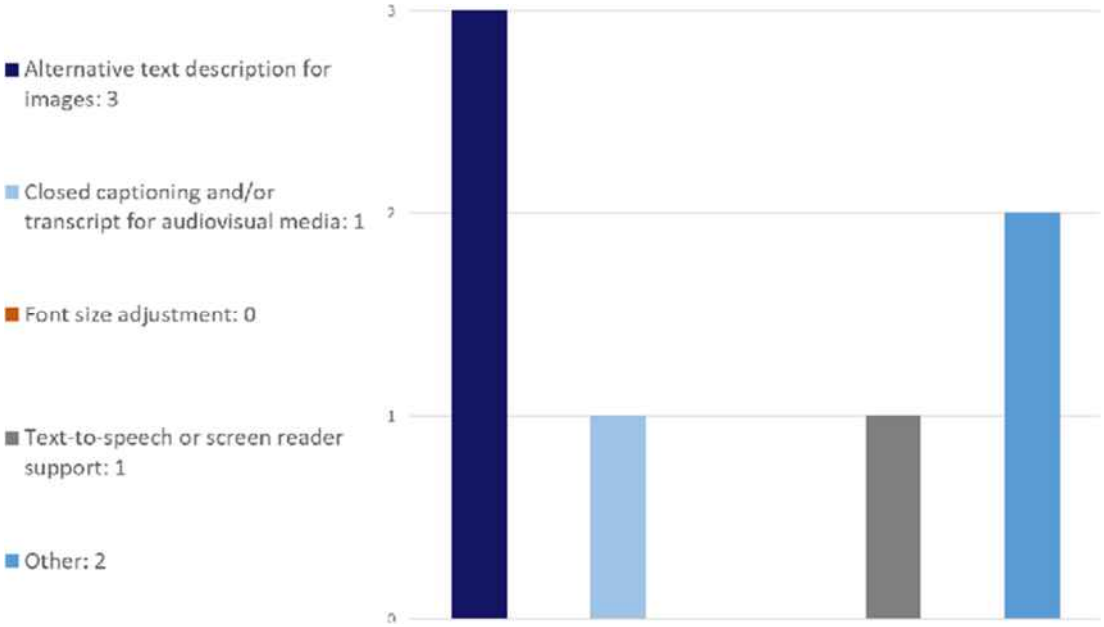
3.4.4 If your organization does not host a repository for research datasets, what, if any, external services are used to provide access?
The named external services used to provide access to research data were Zenodo, GitHub and Radar-Service.

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



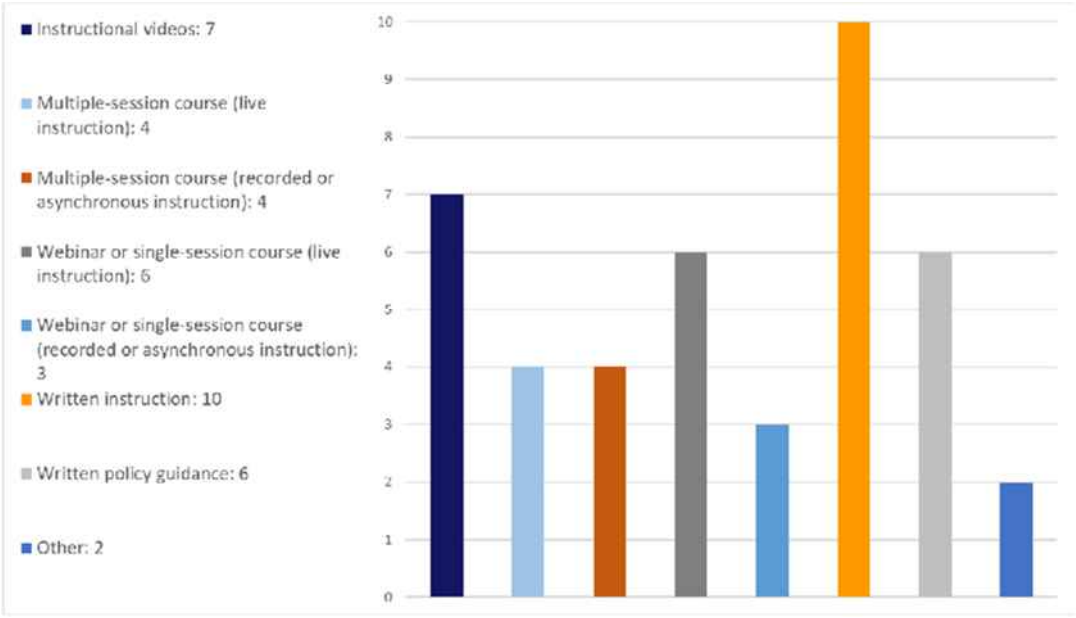
3.4.6 Does your organization provide any of the following infrastructures, tools, or services to make the research datasets repository accessible to users with disabilities?

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



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

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



3.4.8 Does your organization assign persistent identifiers to research datasets?

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

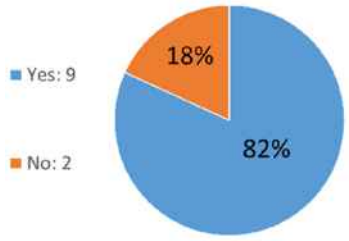
3.4.9 Does your organization assign digital object identifiers (DOIs) to research datasets? Which DOI registration services are used to assign DOIs to research datasets?

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

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

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

Does your organization assign persistent identifiers to research datasets?



Does your organization assign digital object identifiers to research datasets?

- Yes: 7
- No: 2

Which DOI registration services are used to assign DOIs to research datasets?

- DataCite: 5
- Entertainment Identifier Registry (EIDR): 0
- ISTIC & Wanfang Data: 1
- Japan Link Center (JaLC): 1
- Korea Institute of Science and Technology (KISTI): 1
- Multilingual European DOI Registration Agency: 0
- Publications Office of the European Union: 0
- Other: 0

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

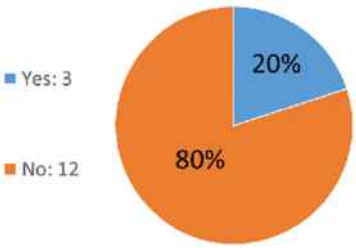
- ARK (archival resource key: 0
- arXiv: 0
- Internal identifier (i.e. organization specific identifier): 3
- URN (uniform resource name): 1
- Other: 2

3.5 Software

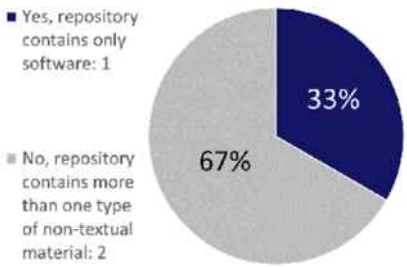
3.5.1 Does your organization host a repository for software? Is the repository exclusive to software? Is the software repository publicly accessible?

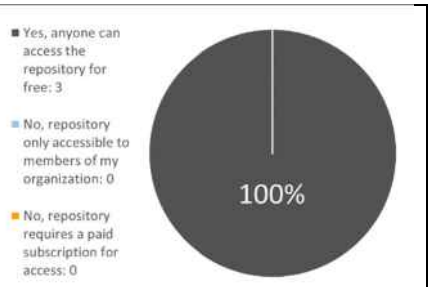
Three of the respondents stated that they have a repository for software, with only one of those being exclusive to containing software only. All three software repositories are publicly available.

Does your organization host a repository for software?



Is the repository exclusive to software?



	Is the software repository publicly accessible?	

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

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

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

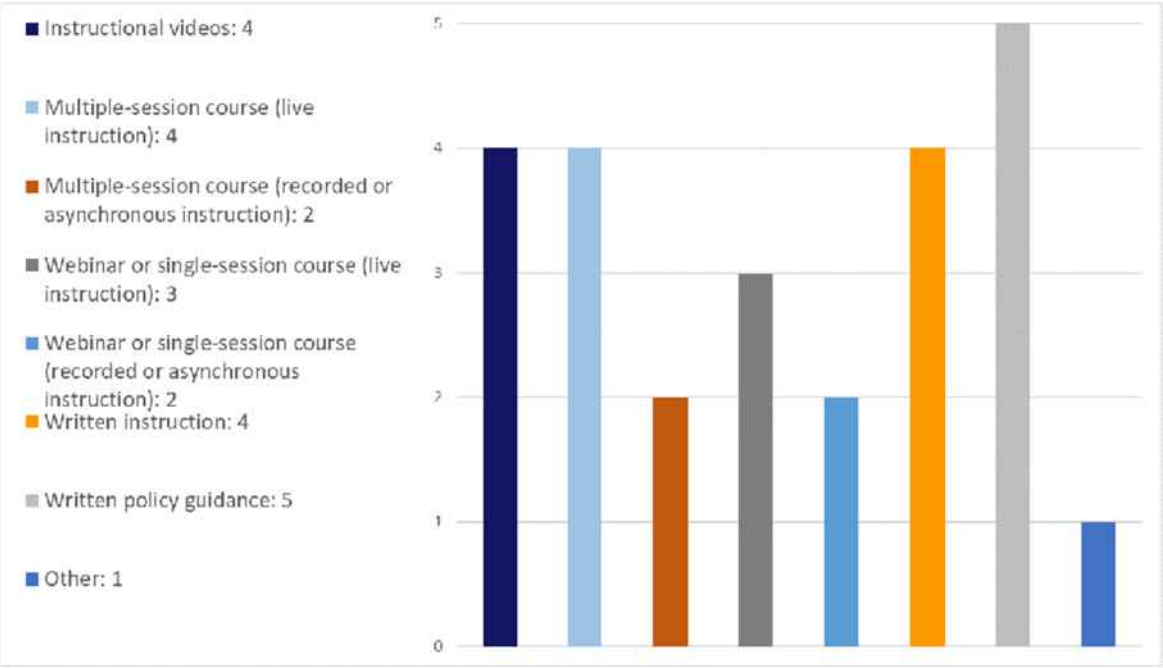
All three of the organizations who host a repository for software provide long-term preservation.

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

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

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

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

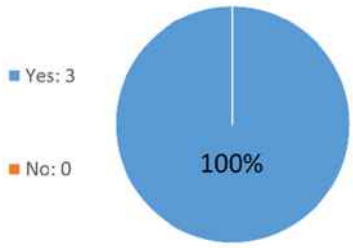


3.5.6 Does your organization assign persistent identifiers to software? Does your organization assign digital object identifiers (DOIs) to software? Which DOI registration services are used to assign DOIs to software?

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

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

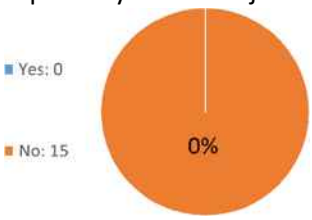
One organization assigns internal identifiers to software.

Does your organization assign persistent identifiers to software?  	Does your organization assign digital object identifiers to software? • Yes: 3 • No: 0 Which DOI registration services are used to assign DOIs to software? • DataCite: 3 • Entertainment Identifier Registry (EIDR): 0 • ISTIC & Wanfang Data: 0 • Japan Link Center (JaLC): 0 • Korea Institute of Science and Technology (KISTI): 0 • Multilingual European DOI Registration Agency: 0 • Publications Office of the European Union: 0 • Other: 0 Does your organization assign any of the following types of persistent identifiers to software (in addition to or instead of DOIs)? • ARK (archival resource key: 0 • arXiv: 0 • Internal identifier (i.e. organization specific identifier): 1 • URN (uniform resource name): 0 • Other: 0

3.6 3D Objects

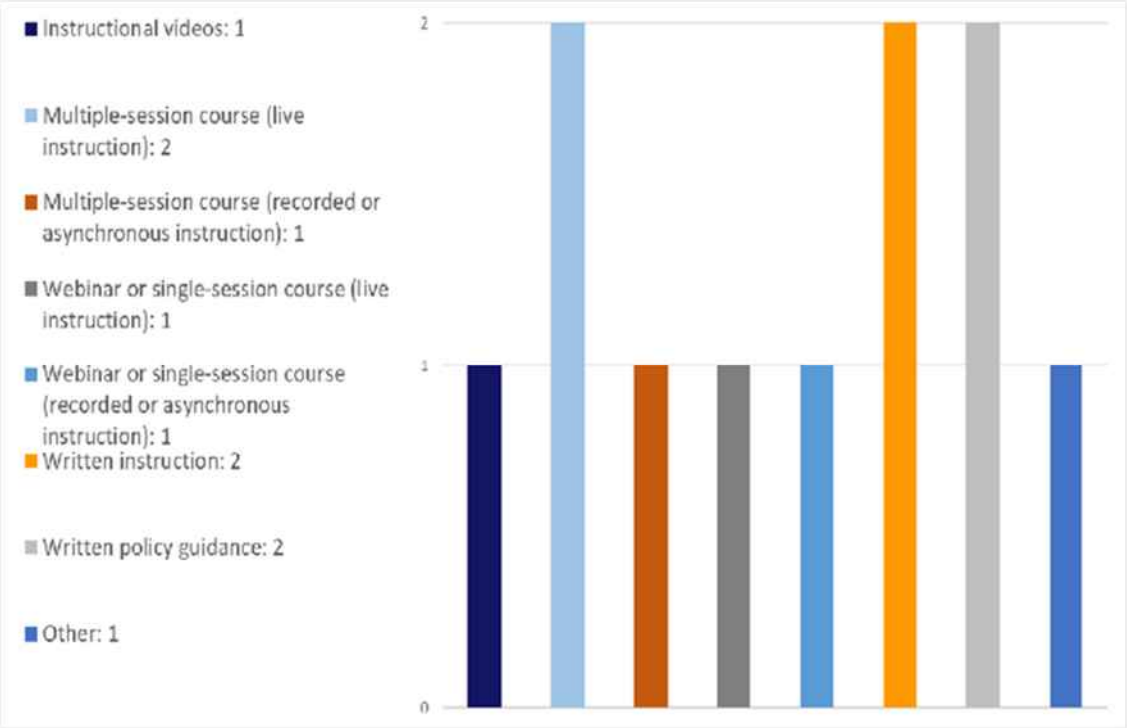
3.6.1 Does your organization host a repository for 3D objects?

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



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

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



3.7 Challenges

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

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

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

4. Summary and Conclusion

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

It appears that research data infrastructures and management are already well established. 75% of the respondents stated that their organization operates its own repository, most of which are publicly available. The vast majority of datasets have a DOI and are archived for the long-term.

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

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

5. Literature

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

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

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

Leadership in Virtual Multicultural Team: A Systematic Literature Review on Definition, Construct, and Future Agenda

Daniel Karim, Aurik Gustomo, and Anggara Wisesa,
School of Business Management, Bandung
Institute of Technology, Indonesia

Abstract
The development on virtual multicultural team studies leads to the concern of leadership's influence to the team. This systematic literature review provides the overview, descriptive analysis, and future suggestions about virtual multicultural team using Scientific Procedures and Rationales for Systematic Literature Reviews (SPAR-4-SLR). This study found four (4) core themes from the literature: CQ, Leadership, team effectiveness, and AI influence. The findings leads to the future suggestions of exploring the role of leadership behaviour and functions in virtual multicultural teams, especially from the soft side of leader-member interaction.

Keywords: *Virtual Team; Multicultural Team; Leadership; CQ*

Introduction

Whenever more than one individual interacts and shares similar characteristics, they are forming a social group. The first conception of Social Group was first introduced by a sociologist Charles Cooley in 1909, in which divide groups into two main levels: primary and secondary. The primary social group consists of kin, family, and close relatives that share implicit things such as love and caring. The secondary social group is proportionately impersonal, transactional, and goal-oriented, such as the working group. However, despite the benefit derived from collective collaboration, there are still debates among scholars about the efficiency and effectiveness of a group that results in the creation of a team which covers individual and collective processes toward common goals[1].

Team practice is evolving from time to time. In modern society, digitalization has become imperative, especially in business organization. Digital transformation disrupts markets and forces traditional firms to change by influencing consumers' expectations and behaviors[2]. The consequences to team practice are adopting digital technology for collaboration and communication who participate in virtual interaction[3], namely virtual team. Moreover, since the virtual activities of a team uncover the possibility of working cross-border, team composition may foster diverse members. Hence the fact that multicultural team topics were developed from a long time before the term "virtual" came to a surface, emerged by the consequences of technology advancement, it adds definitive conception to multicultural teams of being more dynamic and distanced[3]

Multicultural teams by definition are individuals from different cultural backgrounds, bound by social and organizational systems. Leadership of multicultural teams involves collaborating with people from different cultural backgrounds. Leaders in multicultural teams should have a mindset of shared leadership that distributes responsibilities among members[1], leaving a bossy attitude behind. On the other hand, the virtual team also has its own unique perspective in terms of leadership conception that can be seen as a multilevel phenomenon. Leaders in a virtual team may interact with the team, while still engaging personally to each individual[4]. Reflecting the broad study of virtual teams and multicultural teams, we assume that leadership in virtual multicultural teams need to be explored, particularly to construct conceptual models from literature and presenting future research agenda.

This review proposed the current development of virtual multicultural team in reputable journal with high quality paper from Elsevier database from 2018 – 2022, by answering two essential questions of what is the association model for antecedent, moderating, and consequences factors of leadership in virtual multicultural team and which conceptual model facilitates the construct. Therefore, we believe that the update of systematic literature review is helpful to envision the advancement of leadership literature, particularly in the virtual multicultural team.

Method

This study adopts systematic literature review. Referring to well-established methodologies in team management research, this study is using domain-based systematic literature review[4][5][6]. In organizing our literature review, we take on Scientific Procedures and Rationales for Systematic Literature Reviews (SPAR-4-SLR)[7]. With SPAR-4-SLR this study organization includes three stages: assembling, arranging, and assessing the literature.

Assembling

The assembling stages inclusive of the identification and acquisition of literature. In this study, we only use literature from high reputable journals with high quality output from the Elsevier database from 2018-2022 to grasp the trends for today's practices and future references. In creating robust procedures, we make an inclusion and exclusion criteria as follows:

Table 1: Criteria Description

C1	The studies must be the latest in the publication of research, high academic quality standards, and going through peer-review systems. All of the selected journals are from Elsevier and indexed in Scopus, that meet all the standard criteria.
C2	Leadership and Virtual Multicultural Teams context must be central to the manuscript. One of these concepts should at least be part of the title.
C3	The studies must be primary and make explicit or specify the following structure: introduction, methods, results, and discussion (IMRAD structure).

Arranging

Applying these criteria, we conduct several phases of literature search:

1. First search
   - Source : Elsevier
   - Keyword : Virtual Multicultural Teams
   - Type : Research Article
   - Total : 634 articles
   - Subject areas : Business, Management and Accounting (220 articles)
   - Time Frame : 5 years (2018 – 2022), resulting 92 articles
2. Second search
   - Source : Elsevier
   - Keyword : Leadership
   - Type : Research Article
   - Total : 134.251 articles
   - Subject areas : Business, Management and Accounting (28.193 articles)
   - Time Frame : 5 years (2018 – 2022), resulting 8.425 articles

3. Final search
   - Source : Elsevier
   - Keyword : Leadership AND Virtual Multicultural Teams
   - Type : Research Article
   - Total : 246 articles
   - Subject areas : Business, Management and Accounting (105 articles)
   - Time Frame : 5 years (2018 – 2022), resulting 40 articles
   - Purification : 22 articles

Table 2: Final arrangement of literature

No	Journal	Articles
1	Journal of international management	5
2	Journal of business research	4
3	Business Horizon	3
4	International journal of intercultural relations	2
5	Journal of World Business	2
6	The leadership quarterly	1
7	Journal of vocational behavior	1
8	The international journal of management education	2
9	Human resources management review	1
10	Organizational behavior and human decision process	1
	TOTAL	**22**

Special Case: Due to only 22 articles after purification process, the researchers try to include the keyword from relevant subject areas, such as Computer science (4 articles)

No	Journal	Articles
1	Proceedia Computer Science	**4**

Psychology (2 articles)

No	Journal	Articles
1	International journal of clinical and health psychology	1
2	Learning, culture and social interaction	1
	TOTAL	**2**

Resulting a total **of 28 articles** to be included in this systematic literature review analysis.

Assessing

We assessed the literature in three main stages:

1) descriptive analysis to reflect what we know,
2) themes analysis to reflect insight from the literature, and
3) future agenda to know where we are heading next.

DESCRIPTIVE ANALYSIS

Table 3.1: Descriptive characteristics of the study

No	Author(s)	Objective(s)	Finding(s)
1	Allen, et al [8]	Exploring knowledge and skills for HR professional to effectively respond to religious diversity and spirituality at work	The diverse working environment should build trust at cultural intersections by enhancing cultural and religious literacy, fostering interpersonal empathy, and developing inclusive organizational culture.
2	Guler, et al [9]	Using an input-processes-output (IPO) team science model to better understand the team's challenges, limitations, and successes	Composing teams that are demographically diverse on a variety of surface-level (e.g., age, gender, race) and deep-level differences (e.g., discipline, values, abilities) might result in some difficulties in making progress on project tasks. Leadership and its members should handle challenges to achieve goals, including collaborating in setting communication and working standard.
3	Alkhatib, et al [10]	The objective of this paper is to virtually construct three types of teams: innovative, knowledge, and quality based on Herrmann Brain Dominance Instrument (HBDI) profile scores.	Team in low level hierarchical position in organization focuses on silo tasks of process improvement. The middle level team focuses in multidisciplinary integration to strive for effectiveness. And the high level team are responsible to organization and its roots, such as company policy and programme.
4	Zaccaro, et al [11]	To review conceptual framework of how leader attributes influence leadership emergence	High intelligence leader has higher chances to perform better as they could engage in cognitive demanding tasks such as problem solving and divergent thinking. In cross-cultural environment, cultural intelligent and social capacities become fundamental.
5	Webber, et al [12]	Implementation of AI to increase team effectiveness	AI could influence human leader in making decisions. Big data can covers team dynamic and help team leader in developing their members. Unlikely, using AI may cause some challenges for the leaders. They

			need the ability to see micro details. Moreover, AI does not provides personal touch and lack of tolerance. Leaders should balance their role to maintain team motivation and spirit.
6	Caligiuri, et al [13]	Developing cross-cultural competencies through three features of cross cultural experience: contextual novelty, meaningful projects, and social support	Personality plays a significant role in the effectivity of cultural experience. As a collective factors, the three features improve cultural experience, but only contextual novelty has significant impact. The more diverse conditions they experienced, the higher competency improvement.
7	Azevedo, et al [14]	The implementation of pedagogical approach in balancing traditional and experiential approach for a new cultural intelligence training program	Using the new approach, there are increase in innovative work behaviour, resilience and cultural intelligence improvement
8	Popov, et al [15]	The use of interculturally enriched collaboration script (IECS) to enhance student attitudes, behaviour, and learning performance	IECS does improve online collaboration, but not improving performance
9	Demangeot, et al [16]	To actively seeking a well-being-enhancing multicultural engagement	The initiative should be supported by senior leader to adopt in socially-transformative practices
10	Nurhas, et al [17]	investigates the impact of differences in cultural background on perceived technical and operational barriers to iGOAL (Intergenerational collaboration)	The technical barriers in iGOAL are no right time to collaborate, technological complexity, no training for digital collaboration, and difficulty to manage virtual presence. In term of leadership, the focus are in the differences of experience, the subjective perspective of time, and managing virtual presence in working together
11	Super [18]	Presenting the core process of innovative performance,	Leadership must change over time to catch up with the evolution of virtual team. Team and climate play

		foundation of team innovation, and critical leadership for each stage of evolution	an important role in fostering learning cycle. Leadership has two roles during each team formation processes, foundational (trust, climate, creative environment, prosocial motivation, and collective cognition), and progressive (mentor for forming, instructor for storming, coach for norming, and facilitator for performing) from formation to promotion and delegation.
12	Zakaria, et al [19]	Understanding the process of swift trust formation in GVT	Trust as a key factor in virtual multicultural team performance. The dynamic of trust can be viewed as temporal as the virtual team itself. Leader's role is to establish rules, assigning tasks, and monitoring the progress.
13	Cao, et al [20]	Exploring general trust as an important aspect for diverse and temporal team collaboration.	Sense of diversity creates generalization, it helps organization to grasp the feeling of inclusiveness in creating a policy. Then social uncertainty and positive motivation is vital for trust development.
14	Presbitero [21]	Explaining the correlation between CQ and the influence to team effectiveness	Positive correlation of high CQ with high team effectiveness. Leader in GVT can make a direction based on the interpersonal aspects that need to be improved by sharing knowledge and collaboration.
15	Finck, et al [22]	Adaptation and validation of multicultural outcome assessment instrument	Empathy in cultural studies is important to enhance the intercultural competence
16	Stefan, et al [23]	Reviewing Flexible Global Working Engagement to explore a more agile organization and efficient global knowledge transfer	There are lack of HR functions linkage to global talent mobility. Some research has highlighted leadership development in global working environment, but still need to be linked to the integration with HR function in organization.
17	Richter, et al [24]	Correlation of CQ and social integration	Motivational CQ is the most significant factors in social integration.
18	Arslan, et al [25]	Analyze team task environment and	Personal experience and cultural experience are derived from becoming a member of

		personel in multicultural R&D team	multicultural team. It indicates that team task environment is more important than original culture, and leaders should be responsive to increase task variety in a positive manner.
19	Schultheiss [26]	Giving information of women in top position post-pandemic	Moving forward, research needs to incorporate diversity to broaden perspectives, enlighten long-held assumptions, and encourage inclusivity in leaders and leadership styles.
20	Muszynska [27]	Aim to review scientific paper of communication in virtual team	Communication in VT is focused on the tools and methods rather than softer approach such as culture, trust, and leadership
21	Pidduck, et al [28]	Analysing CQ antecedents to understand on how to increase individuals cultural compentencies	Multicultural identity through cross-cultural experience is the antecedents of CQ. More researches are focus on the effectivity of global leadership under CQ studies to translate multicultural experiences into CQ
22	Lacoste, et al [29]	To further the conceptualization of GAM leadership	Expanding the theoretical perspective of global leaders by highlighting managerial roles and behaviours. Vertical and horizontal style of leadership should not be seen as contradictory, rather as complimentary.
23	Gilli, et al [30]	Generalizing teams finding to global virtual team	Increased individual satisfaction with the team's effort and performance is negatively associated with individual performance in GVTs
24	Hundschell, et al [31]	Exploring how subordinates work engagement influenced by the perception of leaders cultural gap bridging behaviour	Leaders CGB behaviour affecting indirectly to subordinates work engagement by giving them sense of multicultural identity
25	Soroka-Potrzebna [32]	Identifying barriers of knowledge management in virtual team	Leaders function in virtual team is to mitigate challenges. Barrier to VT are classified into three level: low level (poor planning and lack of collaboration), middle level

			(operational factor), and high level (IT and motivational tools).
26	Shrivastava [33]	Determining self-report and other report to predict the elimination of cultural conflict	Self-report and other report serves as an information to understand multicultural meaning behind cultural conflict
27	George, et al [34]	Examining traditionalism and virtuality to see the relationship between team effectiveness and shared leadership	Shared leadership has positive impact on team effectiveness. The higher the traditionalism, the higher the need of virtuality.
28	Livermore, et al [35]	The importance of developing organizational CQ	CQ is used for hiring and promoting, learning and development, information sharing, and decision-making. In term of leadership, organizational CQ gives sense of direction in today's diverse, digital world.

Descriptive Analysis Findings

The findings highlight the importance of cultural literacy and empathy by elaborating the significance of enhancing cultural and religious literacy in diverse working environments to build trust at cultural intersections. It emphasizes the role of empathy in fostering interpersonal relationships and creating an inclusive organizational culture. Empathy in cultural studies is seen as important to enhance intercultural competence, and leaders are expected to establish rules, assign tasks, and monitor progress to build trust in virtual multicultural teams.

It also suggest the challenges of demographic and deep-level diversity that composing teams with diverse demographic and deep-level differences may result in difficulties in making progress on project tasks. Leadership and team members are expected to handle these challenges, including setting communication and working standards, to achieve goals in such diverse teams. The role of leadership is highlighted in mitigating challenges and managing differences in virtual teams.

Furthermore, it also highlights the intersection of this study objective about leadership in multicultural and cross-cultural environments. The findings highlight the importance of leadership in multicultural and cross-cultural environments. It suggests that high intelligence leaders may perform better in cognitive demanding tasks, and cultural intelligence and social capacities become fundamental in cross-cultural environments. The use of AI in decision-making is also discussed, with potential challenges and the need for leaders to balance their role to maintain team motivation and spirit.

In supporting leadership in virtual multicultural team, current literature also emphasize the importance of team dynamics and climate. It suggests that team and climate play important roles in fostering learning cycles and that leadership has foundational and progressive roles during team formation processes. Trust is identified as a key factor in virtual multicultural team performance, and leaders are expected to establish rules, assign tasks, and monitor progress to build trust in virtual teams.

Furthermore, the leaders involvement expand the importance of team task environment and task variety. It suggests that team task environment is more important than original culture, and leaders should be responsive to increase task variety in a positive manner. It also emphasizes the need for expanding theoretical perspectives of global leaders by highlighting managerial roles and behaviors, and the role of shared leadership in team effectiveness. This leads to the importance of cultural intelligence (CQ), used for hiring and promoting, learning and development, information sharing, and decision-making. Organizational CQ is seen as giving a sense of direction in today's diverse, digital world. Motivational CQ is identified as the most significant factor in social integration, and there is a positive correlation between high CQ and high team effectiveness.

CQ does not implied the leaders, but also the integration to HR Function. The findings suggest that there is a lack of linkage between HR functions and global talent mobility. It emphasizes the need for leadership development in global working environments to be linked to integration with HR functions in organizations.

Lastly, the most important aspect in multicultural studies is communication. The findings highlight that communication in virtual teams is often focused on tools and methods rather than softer approaches such as culture, trust, and leadership. It suggests the need to broaden perspectives and incorporate diversity in research to encourage inclusivity in leaders and leadership styles.

Overall, the findings emphasize the importance of cultural literacy, empathy, leadership, team dynamics, task environment, and cultural intelligence in diverse working environments. It highlights the challenges and opportunities of working in multicultural and cross-cultural teams, the role of leadership in mitigating challenges and fostering effective virtual teams, and the need for further research to broaden perspectives and encourage inclusivity in leadership. It also emphasizes the need for integration between HR functions and global talent mobility, and the importance of communication and softer approaches in virtual teams.

THEMES ANALYSIS

Nomological Map of Leadership in Virtual Multicultural Team

First, the researchers analyse the literature to create a nomological map using an open coding. There are 40 codes that are relevant to the study, as shown in the table below.

Table 3.2: Nomological map using an open coding

Codes 1-10	Codes 11-20	Codes 21-30	Codes 31-40
1) Cultural and religious literacy 2) Interpersonal empathy 3) Inclusive organizational culture 4) Demographic and deep-level diversity in teams 5) Leadership 6) Team collaboration	11) AI's influence on human leaders 12) Decision-making in AI-driven teams 13) Personality and cultural experience 14) Contextual novelty 15) Innovative work behavior 16) Resilience 17) Online collaboration	21) Trust in virtual teams 22) Sense of diversity 23) Positive motivation 24) Team effectiveness 25) Empathy in intercultural competence 26) HR functions and global talent mobility 27) Motivational CQ	31) Cross-cultural experience 32) Expanding theoretical perspective of global leaders 33) Individual satisfaction with team effort/performanc e 34) Leaders' CGB behavior

7) Low, middle, and high-level teams 8) High intelligence leader 9) Cultural intelligence 10) Social capacities	18) Senior leader support 19) Technical barriers in virtual teams 20) Team formation processes	28) Team task environment 29) Communication in virtual teams 30) Multicultural identity	35) Function of leaders in virtual teams 36) Self-report and other report in understanding cultural conflict 37) Shared leadership 38) Traditionalism and virtuality 39) CQ's application in various organizational processes 40) Organizational CQ for direction in a diverse, digital world.

Then, the researchers are building a connection between variables to create a relevant model from previous studies. Based on the open coding of the identified themes, here are four core themes that can serve as umbrella categories for all the codes and themes:

1. Cultural Intelligence (CQ)

   This core theme encompasses themes such as cultural and religious literacy [8], interpersonal empathy, demographic and deep-level diversity in teams [9], cultural intelligence, social capacities, personality and cultural experience, contextual novelty, multicultural identity, cross-cultural experience, expanding theoretical perspective of global leaders, and organizational CQ for direction in a diverse [21], digital world. It relates to the ability of individuals and teams to effectively understand, adapt to, and navigate cultural differences and diversity [28] in various contexts, including virtual teams and AI-driven teams.

2. Leadership

   This core theme encompasses themes such as leadership, high intelligence leader, team collaboration, leadership functions in virtual teams, senior leader support, team formation processes, trust in virtual teams, leaders' CGB (Cultural Generativity Behavior) behaviour [31], function of leaders in virtual teams, shared leadership, and traditionalism and virtuality. It relates to the role of leaders in managing diverse teams, fostering inclusive organizational cultures, and promoting effective team collaboration [11], especially in the context of virtual teams and AI-driven teams [12].

3. Team Effectiveness

   This core theme encompasses themes such as team effectiveness, low, middle, and high-level teams, innovative work behavior, resilience, online collaboration, team task environment, communication in virtual teams [27], individual satisfaction with team effort/performance, and self-report and other report in understanding cultural conflict [33]. It relates to the factors that contribute to or hinder the effectiveness of teams, including virtual teams, in diverse and multicultural contexts, and the role of team dynamics, communication, and collaboration in achieving positive outcomes.

4. AI's Influence

   This core theme encompasses themes such as AI's influence on human leaders, decision-making in AI-driven teams, and technical barriers in virtual teams. It relates to the impact of artificial intelligence (AI) on leadership, team dynamics, and organizational processes, including decision-making, communication, and collaboration in virtual teams, and the challenges and opportunities posed by AI in diverse and multicultural contexts [12].

These four core themes can provide a conceptual framework to understand the relationships and interactions among the various codes and themes identified in the findings, and can serve as a foundation for further analysis and interpretation of the data.

Conclusion

In conclusion, this systematic literature review has provided insights into the topic of leadership in virtual multicultural teams, focusing on the definition, construct, and future agenda. Through an in-depth analysis of the literature, several key themes and codes have emerged, which can be categorized into four core themes: Cultural Intelligence (CQ), Leadership, Team Effectiveness, and AI's Influence.

The core theme of Cultural Intelligence (CQ) highlights the importance of understanding, adapting to, and navigating cultural differences and diversity in virtual multicultural teams. This includes themes such as cultural and religious literacy [8], interpersonal empathy, and cross-cultural experience, which are crucial for leaders to effectively manage multicultural teams in virtual environments.

The core theme of Leadership underscores the critical role of leaders in virtual multicultural teams [16], including their functions, behaviors, and support. This includes themes such as leadership, team collaboration, shared leadership, and traditionalism, which highlight the importance of inclusive leadership and fostering collaborative team dynamics in virtual multicultural teams.

The core theme of Team Effectiveness emphasizes the factors that contribute to or hinder the effectiveness of virtual multicultural teams. This includes themes such as team effectiveness, innovative work behavior, and resilience, which emphasize the significance of team dynamics, communication, and collaboration in achieving positive outcomes in multicultural virtual teams.

The core theme of AI's Influence highlights the impact of artificial intelligence on leadership, team dynamics, and organizational processes in virtual multicultural teams. This includes themes such as AI's influence on human leaders and decision-making in AI-driven teams, which highlight the challenges and opportunities posed by AI in diverse and multicultural contexts.

Overall, this systematic literature review provides a comprehensive understanding of the key concepts, constructs, and future agenda related to leadership in virtual multicultural teams. The identified core themes can serve as a foundation for further research and practical implications in the field of leadership in virtual multicultural teams, informing future studies and interventions aimed at enhancing leadership effectiveness, team dynamics, and overall team performance in diverse virtual environments.

FUTURE AGENDA

Gap, Limitations, and Future Research Suggestions

Here are the potential gaps, limitations, and future research agenda that could be explored:

1. Gap

   First, limited research on the impact of AI on virtual multicultural teams. While the model identifies the influence of AI on leadership and team dynamics, there may be limited research on the micro details, personal touch, and team motivation aspects of AI's influence in virtual multicultural teams. Second, limited research on the link between cultural intelligence (CQ) and social capacities. While the model identifies cultural intelligence and social capacities as core themes, there may be limited research on the relationship between these variables and how they interact in virtual multicultural teams.

2. Future Research Agenda

   First, exploring the role of leadership behaviors and functions in virtual multicultural teams. Further research could focus on identifying specific leadership behaviors and functions that are most effective in virtual multicultural teams, such as setting communication and working standards, handling challenges, and achieving team goals. Second, Examining the impact of team diversity on team effectiveness. Further research could investigate the relationship between demographic and deep-level diversity in virtual multicultural teams and their impact on team effectiveness, including knowledge sharing, collaboration, and trust [19]. Third, integrating HR functions with leadership development in virtual multicultural teams [8]. Future research could explore the linkage between HR functions, such as global talent mobility, and leadership development in virtual multicultural teams to enhance organizational effectiveness and inclusivity. Lastly, exploring the impact of cultural identity and cross-cultural experience on leadership effectiveness. Further research could investigate how multicultural identity and cross-cultural experience of leaders influence their leadership effectiveness in virtual multicultural teams, and how these factors may interact with other variables such as cultural intelligence and social capacities.

Overall, there is a need for more empirical research to further understand the complex dynamics of leadership in virtual multicultural teams and to explore new directions for future research to address the identified gaps and limitations in the literature.

Glossary

Term	**Definition**
AI (Artificial Intelligence)	A reproduction of human being cognitive abilities by a machine
Contextual Novelty	Unexpected event from a particular context
CQ (Cultural Intelligence)	A set of ability to understand diversity of culture
Cultural Literacy	The ability to understand tradition of people from particular culture
Deep Level Diversity	Non-observable characteristics such as belief and values
Inclusive Organizational Culture	A culture of an organization that respect every belief, values, and culture at every level
Social Capacities	An ability to adapt in social situation
Traditionalism	Following a doctrine or the maintaining behavior of a tradition
Virtuality	A consideration or product of imagination from a fact as a reflection of a real thing

References

[1] S. A. Halverson, Claire B; Tirmizi, *Effective Multicultural Teams: Theory and Practice*. 2008.

[2] P. C. Verhoef *et al.*, "Digital transformation: A multidisciplinary reflection and research agenda," *J. Bus. Res.*, vol. 122, no. November 2019, pp. 889–901, 2021.

[3] K. Cagiltay, B. Bichelmeyer, and G. Kaplan Akilli, "Working with multicultural virtual teams: critical factors for facilitation, satisfaction and success," *Smart Learn. Environ.*, vol. 2, no. 1, 2015.

[4] C. Liao, "Leadership in virtual teams: A multilevel perspective," *Hum. Resour. Manag. Rev.*, vol. 27, no. 4, pp. 648–659, 2017.

[5] D. Tranfield, D. Denyer, and P. Smart, "Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review* Introduction: the need for an evidence- informed approach," *Br. J. Manag.*, vol. 14, pp. 207–222, 2003.

[6] M. Russen, M. Dawson, and J. M. Madera, "Gender diversity in hospitality and tourism top management teams: A systematic review of the last 10 years," *Int. J. Hosp. Manag.*, vol. 95, no. March, p. 102942, 2021.

[7] J. Paul, W. M. Lim, A. O'Cass, A. W. Hao, and S. Bresciani, "Scientific procedures and rationales for systematic literature reviews (SPAR-4-SLR)," *Int. J. Consum. Stud.*, no. April, pp. 1–16, 2021.

[8] S. Allen, P. Williams, and D. Allen, "Human resource professionals' competencies for pluralistic workplaces," *Int. J. Manag. Educ.*, vol. 16, no. 2, pp. 309–320, 2018.

[9] J. Guler *et al.*, "Global Collaborative Team Performance for the Revision of the International Classification of Diseases: A Case Study of the World Health Organization Field Studies Coordination Group," *Int. J. Clin. Heal. Psychol.*, vol. 18, no. 3, pp. 189–200, 2018.

[10] G. Alkhatib and O. Al-Humaidi, "Innovative virtual teams on demand: HBDI-based paradigm," *Procedia Comput. Sci.*, vol. 131, pp. 139–147, 2018.

[11] S. J. Zaccaro, J. P. Green, S. Dubrow, and M. J. Kolze, "Leader individual differences, situational parameters, and leadership outcomes: A comprehensive review and integration," *Leadersh. Q.*, vol. 29, no. 1, pp. 2–43, 2018.

[12] S. S. Webber, J. Detjen, T. L. MacLean, and D. Thomas, "Team challenges: Is artificial intelligence the solution?," *Bus. Horiz.*, vol. 62, no. 6, pp. 741–750, 2019.

[13] P. Caligiuri, A. Mencin, B. Jayne, and A. Traylor, "Developing cross-cultural competencies through international corporate volunteerism," *J. World Bus.*, vol. 54, no. 1, pp. 14–23, 2019.

[14] A. Azevedo and M. J. Shane, "A new training program in developing cultural intelligence can also improve innovative work behavior and resilience: A longitudinal pilot study of graduate students and professional employees," *Int. J. Manag. Educ.*, vol. 17, no. 3, p. 100303, 2019.

[15] V. Popov *et al.*, "Effects of an interculturally enriched collaboration script on student attitudes, behavior, and learning performance in a CSCL environment," *Learn. Cult. Soc. Interact.*, vol. 21, no. August 2018, pp. 100–123, 2019.

[16] C. Demangeot *et al.*, "Constructing a bridge to multicultural marketplace well-being: A consumer-centered framework for marketer action," *J. Bus. Res.*, vol. 100, no. December 2017, pp. 339–353, 2019.

[17] I. Nurhas, B. R. Aditya, S. Geisler, and J. Pawlowski, "Why does cultural diversity foster technology-enabled intergenerational collaboration?," *Procedia Comput. Sci.*, vol. 161, pp. 15–22, 2019.

[18] J. F. Super, "Building innovative teams: Leadership strategies across the various stages of team development," *Bus. Horiz.*, vol. 63, no. 4, pp. 553–563, 2020.

[19] N. Zakaria and S. A. Mohd Yusof, "Crossing Cultural Boundaries Using the Internet: Toward Building a Model of Swift Trust Formation in Global Virtual Teams," *J. Int. Manag.*, vol. 26, no. 1, p. 100654, 2020.

[20] J. Cao and A. D. Galinsky, "The Diversity-Uncertainty-Valence (DUV) model of generalized trust development," *Organ. Behav. Hum. Decis. Process.*, vol. 161, no. March, pp. 49–64, 2020.

[21] A. Presbitero, "Communication accommodation within global virtual team: The influence of cultural intelligence and the impact on interpersonal process effectiveness," *J. Int. Manag.*, vol. 27, no. 1, p. 100809, 2021.

[22] C. Finck, Y. Gómez, J. N. Castro, E. Y. Mogollón, N. Marcelo, and A. Hinz, "Adaptation and validation of a Spanish instrument for assessing multicultural competencies and empathy," *Int. J. Intercult. Relations*, vol. 83, no. June, pp. 163–176, 2021.

[23] S. Jooss, A. McDonnell, and K. Conroy, "Flexible global working arrangements: An integrative review and future research agenda," *Hum. Resour. Manag. Rev.*, vol. 31, no. 4, p. 100780, 2021.

[24] N. F. Richter, J. Martin, S. V. Hansen, V. Taras, and I. Alon, "Motivational configurations of cultural intelligence, social integration, and performance in global virtual teams," *J. Bus. Res.*, vol. 129, no. March 2020, pp. 351–367, 2021.

[25] A. Arslan, L. Haapanen, P. Ahokangas, and S. Naughton, "Multicultural R&D team operations in high-tech SMEs: Role of team task environment and individual team members' personal experiences," *J. Bus. Res.*, vol. 128, no. January 2020, pp. 661–672, 2021.

[26] D. E. Schultheiss, "Shining the light on women's work, this time brighter: Let's start at the top," *J. Vocat. Behav.*, vol. 126, no. April, p. 103558, 2021.

[27] K. Muszynska, "A bibliometric review of research on communication in virtual project teams," *Procedia Comput. Sci.*, vol. 192, pp. 4770–4779, 2021.

[28] R. J. Pidduck, M. A. Shaffer, Y. Zhang, S. S. Y. Cheung, and D. G. Yunlu, "Cultural intelligence: An identity lens on the influence of cross-cultural experience," *J. Int. Manag.*, vol. 28, no. 3, p. 100928, 2022.

[29] S. Lacoste, K. Zidani, and J. M. Cuevas, "Lateral collaboration and boundary-spanning from a global leadership perspective: The case of global account managers," *J. World Bus.*, vol. 57, no. 3, p. 101288, 2022.

[30] K. Gilli, V. Veglio, M. Gunkel, and V. Taras, "In search of the Holy Grail in global virtual teams: The mediating role of satisfaction on performance outcomes," *J. Bus. Res.*, vol. 146, no. March, pp. 325–337, 2022.

[31] A. Hundschell, J. Backmann, A. W. Tian, and M. Hoegl, "Leaders' cultural gap bridging behaviors and subordinates' work engagement in multinational teams," *J. Int. Manag.*, vol. 28, no. 3, p. 100916, 2022.

[32] H. Soroka-Potrzebna, "Barriers of knowledge management in virtual project teams: a TISM model," *Procedia Comput. Sci.*, vol. 207, pp. 800–809, 2022.

[33] A. Shrivastava and A. Pandey, "Understanding intra-group response biases in culturally diverse teams," *Int. J. Intercult. Relations*, vol. 88, no. April, pp. 189–199, 2022.

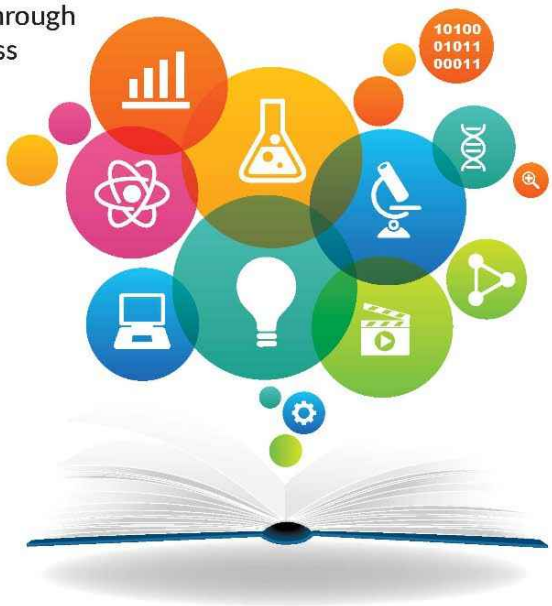
[34] C. George, C. B. Gibson, and J. Barbour, "Shared leadership across cultures: Do traditionalism and virtuality matter?," *J. Int. Manag.*, vol. 28, no. 1, p. 100905, 2022.

[35] D. Livermore, L. Van Dyne, and S. Ang, "Organizational CQ: Cultural intelligence for 21st-century organizations," *Bus. Horiz.*, vol. 65, no. 5, pp. 671–680, 2022.

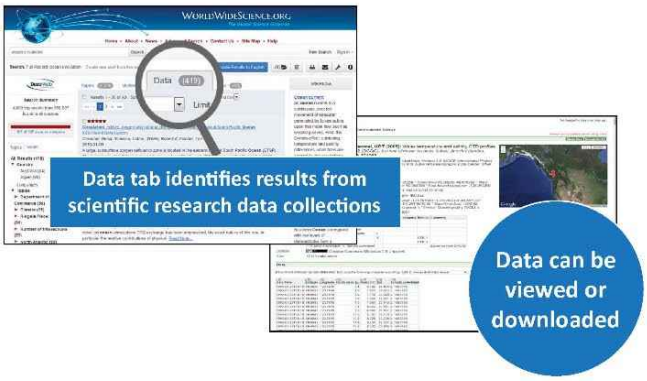
WorldWideScience.org

An International Partnership Supporting Open Science

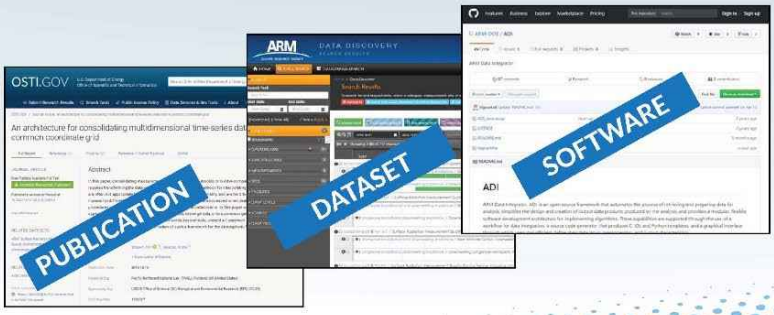
- Simultaneously explore over 100 national and international scientific databases and portals from 75 countries
- Search information in textual, multimedia, software, and scientific data formats
- Eliminate language barriers through multilingual translations across ten languages



Search and Access Scientific Research Data



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



WorldWide ScienceAlliance

Operating Agent:

OSTI.GOV

WorldWideScience.org Operating Agent
Point of Contact: Lorrie Johnson, JohnsonL@osti.gov

Cheap, Rigorous, and Transparent: How Web-scraping with Python can Improve Collecting Grey Literature for Systematic Literature Reviews

Cameron Atkinson, School of Social Sciences, University of Tasmania, Australia

Abstract

Gathering non-conventional literature, such as grey literature, from web-based sources for use in a systematic literature review is at times an arduous task. Often, the processes used to do so are difficult for other researchers to repeat. Compounding this issue is the cost that researchers bear, in either paying for desktop-based applications, or paying external researchers who have programming experience to design applications and tools for them. To address these issues, this article presents a methodology for researchers to systematically gather grey literature from online repositories using the computer programming language Python. Utilising a well-known data extraction technique (Web-Scraping), this article exhibits the code used to scrape policy documents from the International Energy Agency's online policy database. A flowchart for the different stages of this process is also introduced to aid in addressing the technical, legal, and ethical elements of web-scraping that researchers must also be aware of when undertaking this approach. Finally, as a proof of concept, the results from the method are also presented.

Keywords: *Systematic Literature Review, Grey Literature, Computer Programming, Python*

Introduction

Grey literature is one of the many types of grey information. Introduced by Adams et al. (2016) grey information is utilised to distinguish the different forms of 'grey'; including data, literature, and information. Grey 'data' can refer to user-generated web content such as tweets or blogs, 'literature' can refer to policy documents or standards and regulations, and 'information' can be informal information such as meeting notes or emails (Banks, 2009; Kamei et al., 2020). The focus of this article will be on 'grey literature', specifically policy documents held in web-based repositories.

Non-conventional or grey literature is routinely one of the most difficult literatures to methodically gather for use in a systematic review (Paez, 2017). The difficulties in systematically gathering grey literature can lead to its ad hoc inclusion, or even the complete sidelining of it from a review. In both cases, the ignoring of, or ad hoc inclusion of, grey literature becomes a publication bias within the review (Haddaway et al., 2020). To avoid grey literature publication biases, researchers can spend months trying to gather, appraise, and then synthesise grey literature, often through processes that are not repeatable and introduce further biases (Adams et al., 2016). Indeed, the issues faced with replicability of research is part of an ongoing methodological problem where it is difficult, if not impossible to replicate the results of many scientific projects (Simova, 2023). To address the replicability issues surrounding the gathering of grey literature, this article proposes a well-known data extraction technique; web-scraping.

Web-scraping through Application Programming Interfaces (APIs) and computer programming-based tools have been the norm to gather web-based data. However, recent changes have occurred in the way that companies allow researchers access to their APIs. This can be seen with respect to Facebook in 2018 and Twitter 2023, both of whom are limiting access to APIs or charging for access (Anderljung & Hazell, 2023; Freelon, 2018). These changes serve to highlight the importance of programming-based web-scraping tools. There are many commercially available web-scraping programs that will help researchers identify grey literature. However, early doctoral candidates and early career researchers can face financial constraints that effect their ability to purchase program

subscriptions (Caretta et al., 2018). Additionally, as the internet is not a homogonous collection of websites, 'point and click' desktop software may not be able to address the nuances of different webpage designs (Krotov et al., 2020). To address these problems, this paper seeks to highlight a free, rigorous, and transparent methodology for researchers to systematically gather grey literature from web-based repositories. Utilizing the computer programming language Python, this article sets out the steps taken to create a web-scraping tool employed in scraping policy documents from the International Energy Agency's policy repository.

This article presents a method for researchers with some coding experience. As the web-scraper was specifically built for scraping the International Energy Agency's policy repository, this article should be taken as a proof of concept. Furthermore, it illustrates the value that researchers can derive from employing techniques, such as web-scraping, to automate tasks that would otherwise be time consuming, expensive, and difficult for external researchers to repeat. Additionally, as many companies are now shutting down access to their APIs; Facebook (now Meta) in 2018, Twitter and Reddit in 2023 (Anderljung & Hazell, 2023; Freelon, 2018), web-scraping is useful skill for researchers to have (Freelon, 2018). Indeed, Freelon (2018) warned that, although APIs are easy to use and compliant with websites terms of service, much like Facebook in 2018, they can be taken away with the figurative flick of a switch, as we are now seeing with Twitter charging for API access, and Reddit seeking to also monetise their vast data repositories. To the authors knowledge, this is the first time an article has in conjunction discussed the importance of web-scraping for increasing the rigor and repeatability of gathering grey literature, while also providing the code and an example of scraping grey literature from a web-based repository.

The following background section will provide an overview of each element in this article. Firstly, it will briefly introduce the reader to the current crisis surrounding replication in science. It will then continue with another brief introduction on systematic literature reviews and the importance of including grey literature in them. It will then discuss the concept of web-scraping, both API and programming based, and how web-scraping can be applied towards grey literature. Following this, the background section will introduce the computer programming language that is used in this article, the libraries utilised in constructing the code for the proposed web-scraping tool, and the ethical and legal elements that must be kept in mind when utilising the proposed methodology. Next, the method section will provide a complete account for the coding utilised in building the web-scraper before the result section provides methods to view results within the Jupyter Notebook it was written in, and how to save a CSV file of the data scraped for further processing or analysis. Finally, the discussion and conclusion section will summarise the method and the value that this method, and variants of it, will have for researchers conducting systematic literature reviews that include grey literature.

Background

The replicability crisis

Replicability is broadly accepted to provide the foundation for the epistemic authority of science (Romero, 2019). In the previous decade, publications in both scientific journals as well as the press have drawn attention to what has been termed a 'replicability crisis' within research (Shrout & Rodgers, 2018). This crisis is one that has engulfed not just the natural sciences, but the social sciences as well (Dreber & Johannesson, 2019; Hensel, 2021). At its core, it refers to the inability of external researchers to replicate or reproduce results from important findings (Romero, 2019). It has led to the retraction of many research articles due to an inability to reproduce findings (Silva & Nazarovets, 2023).To address the crisis, researchers have been adapting methodologies to include

the publishing of raw data sets, codes employed, alongside more detailed analyses, all under the paradigm of 'open science' (Crüwell et al., 2019; Echtler & Häußler, 2018; Shmagun et al., 2022). One methodology for enhanced integration with the open science mantra of 'Open Methodology, Open Data, Open Source and Open Access' is the systematic review (Haddaway, 2018).

Systematic literature reviews

Literature reviews play an crucial function in academic research to gather existing expertise and to investigate the state of a research domain (Linnenluecke et al., 2020). There are many ways that researchers can use to review the literature in their field; traditional, structured, narrative, and systematic for example (Stratton, 2019). What sets systematic literature reviews apart from other methods of review is the replicable, scientific and transparent process adopted by reviewers (Kitchenham et al., 2009). This approach aims to minimise bias by extensively searching through the literature, both published and unpublished, and providing an evidence trail so that decisions, methods, and conclusions can be questioned (Tranfield et al., 2003). Indeed, systematic review methods are being infused with open science to ensure increased transparency, replicability, and robustness, of the systematic review process (Gunnell et al., 2022). As many findings from systematic literature reviews are used as a basis for evidence informed policy, replicability and transparency of results are a key factor. For researchers aiming to utilise systematic reviews to inform policy, open science not only means access to, and the reporting of, methods towards gathering published research findings, but also to grey literature (Sheehan, 2023).

Grey Literature and Information

The term 'grey literature' became common research parlance in the 1970s (Bogdanski et al., 2005). Traditionally, grey literature was considered to be unpublished research (Benzies et al., 2006). However, according to Sharif et al. (2023) the most extensively used definition for grey literature is,

> *'document types produced on all levels of government, academics, business and industry in print and electronic formats that are protected by intellectual property rights, of sufficient quality to be collected and preserved by library holdings or institutional repositories.'*

Grey literature is so-termed as much, if not all, of its materials do not have an international standard book number (ISBN) or international standard serial number (ISSN), hindering its access it without extra effort (Lawal et al., 2023). Much contemporary grey literature is to be now sourced online (Benzies et al., 2006). Indeed, alongside Google, there are even online databases that are dedicated to grey literature entirely, some including academic literature as well (Mahood et al., 2014). One useful methodology for gathering grey literature for a systematic literature review can be found in Godin et al. (2015). In this they set out a structured means to gather grey literature through; (1) grey literature databases; (2) tailored Google search engines; (3) websites; and (4) contacting experts (Godin et al., 2015). Other guides to gather grey literature for reviews can be found in (Blackhall, 2007; Kugley et al., 2016). Although these guides and research projects highlight the usefulness of Google searches, databases, and websites, they do not discuss the usefulness of web-scraping in gathering grey literature for a systematic literature review. Particularly, how programming-based web-scrapers can aid in gathering grey literature for a systematic literature review.

Web-scraping: APIs and programming-based tools.

The internet is a gigantic source of information for a variety of disciplines, practices, and fields. Manually clicking, opening, viewing, scanning, and assessing documents is a time consuming, and

often unstructured way to gather web-based information. However, there are methods available to make this a more rigorous and structured approach.

Web-scraping, also known as 'Web Harvesting', 'Web Data Extraction' and 'Screen Scraping', is defined as, 'a procedure of automatic web data extraction instead of manually copying it. It is a technique in which meaningful data from the HTML of websites are extracted and stored into a central local database or spreadsheet' (Hannes, 2023). A simple and popular way of achieving this is through accessing information via a websites API.

An API is an interface of a computer program that permits software to communicate with other software, and they are made accessible to third parties by companies (Lomborg & Bechmann, 2014). APIs provided by websites began to appear in the early 2000s, the first of which was by salesforce.com (Krijnen et al., 2014). APIs were developed for the specific purpose of making it simpler and more reliable to access and gather web-based information (Krijnen et al., 2014). However, many companies are now shutting down or charging for access to their APIs; Facebook (now Meta) in 2018 and Twitter in 2023, meaning that web-scraping using programming-based tools has become a useful skill to possess (Freelon, 2018). Indeed, Freelon (2018) warned that, although APIs are easy to use and compliant with a websites terms of service, much like Facebook in 2018, they can be taken away with the figurative flick of a switch. In a similar vein, companies who understand the value of the data in their possession are now charging for API access, as we are now seeing with Twitter and Reddit in 2023 (Anderljung & Hazell, 2023). This leads to an argument that traditional approaches to gathering web-based data (via programming-based scraping tools) are useful skills for researchers to possess.

A web-scraper is a software program that mimics human internet browsing to gather specific data from different websites (Diouf et al., 2019). Data scraping can be achieved in many ways; via a proprietary software, external custom web-scraping services, or via open-source software like Python or R (Luscombe et al., 2022). For propriety software, there have been a number of software packages developed to facilitate the scraping of web-based content; NVivo's NCapture, Tweetreach, Helium Scraper, Qualtrics, Screen Scrapper, OutWit Hub, FMiner, Kimono for example (Haddaway, 2015; Luscombe et al., 2022).

A problem with desktop-based or 'point and click' approaches is that they often do not work as expected. They either miss data, or are simply not smart enough to determine how to access the required data points from a website (Krotov et al., 2020). Indeed, it has been pointed out that when compared to free and open-source programming languages such as Python, propriety software (for example Nvivo's NCapture) can make it difficult for future researchers to reproduce or repeat findings (Luscombe et al., 2022). Furthermore, that software such as NCapture are not only costly, but are not as transparent as programming-based methods (Luscombe et al., 2022). In addition, due to the fluidity of the internet, even advanced web-scrapping software programs may still need custom programming code to ensure that they work as intended (Krotov et al., 2020).

Aside from desktop based software, one of the most common ways to develop web-scraping applications is through the use of third party libraries installed and run with computer programming languages (Glez-Peña et al., 2014). Using the programming language that the researcher is most familiar with, third party libraries, such as the ones available with Python, allow researchers access to websites for scraping by making HTTP requests to the web servers hosting those sites. The next section will discuss how the computer programming language Python can be used to scrape grey literature from websites.

Web-scraping with Python

All programming languages are comprised of code. Codes are sets of written instructions that tell the computer to perform certain tasks (Blackwell, 2002). One of many programming languages, Python is touted as easy to learn, freely available, and an ideal language for scripting and application development (Van Rossum & Drake Jr, 1995). Building web-scraping tools with Python requires utilising different libraries depending on the tools requirements. For this article BeautifulSoup and Requests were utilised to create the web-scraper. The library 'Tabulate' was used to display the data scraped within the Jupiter Notebook the tool was coded in, and the library 'CSV' was employed to save the information scraped into an easily accessible Excel file. There are many introductory resources available for researchers who wish to develop programming skills with Python (please see (Gupta & Sehgal, 2021; Matthes, 2023; Vasiliev, 2020) for further reading). As this article is about web-scraping, only the libraries pertaining to building the web-scraper will be next introduced. However, if the reader is interested in reading about 'Tabulate' and 'CSV' then please see Astanin (2022) for 'Tabulate' and The Python Standard Library (2023) for information regarding 'CSV.'

BeautifulSoup. *BeautifulSoup* is a Python library created by Leonard Richardson among other developers (Uzun et al., 2018). It is primarily intended to parse and extract data from an Hypertext Markup Language (HTML) string (Hajba & Hajba, 2018). HTML is a markup language used for creating the structure and content of web pages (Willard, 2003).

There are many guides available for researchers wanting to learn to web-scrape with *BeautifulSoup,* for example see (Hajba & Hajba, 2018; Nair, 2014). In the code presented in the method section, *BeautifulSoup* is used for parsing the HTML content pulled into the program via *Requests.* Since the text gathered by *Requests* is just a long string, *BeautifulSoup* helps in translating this raw text into a navigable Python object that allows the user to search for specific elements, extract data, and easily navigate through the HTML data.

Requests. To use the internet, the web browser on a computer must communicate with a web server. Whenever a browser communicates with a server, it is done using Hypertext Transfer Protocol (HTTP) which acts as a request-response protocol (Chandra & Varanasi, 2015). HTTP define the sets of rules and conventions for requesting and transmitting data over the internet (Berners-Lee et al., 1996).

For the code used in this method, *Requests* is used to get the content from the IEA policy database via its Uniform Resource Location (URL). The response contains the raw HTML content off the web page that is then parsed through *BeautifulSoup.* As with *BeautifulSoup,* there are many guides available for those wishing to learn how to use *Requests* (please see Chandra and Varanasi (2015) as an example).

Problems with programming-based scraping

One of the largest differences between scraping with an API and scraping with a programming-based tool is speed; APIs are generally faster (Dongo et al., 2020). Another problem is that some websites do not allow for scraping. Furthermore, should a website developer change the architecture of the site being scrapped then the tool may no longer function as intended (Dongo et al., 2020). Technical problems aside, the legal and ethical elements of web-scraping are an important consideration that researchers must be aware of. In the past there have been several legal cases involving the scraping of web pages. Such cases include, but are not limited to prohibited access and use of data, contract breach, and copyright infringement (Krotov et al., 2020).

However, there are some simple ways to ensure ethical and legal factors are considered. For example, permission from websites can be sought prior to scraping. Additionally, researchers (as is the case with the presented method below) can add details of contact and intention by using the User-Agent request header (Mazilu, 2022). The User-Agent request header is, 'a characteristic string that lets servers and network peers identify the application, operating system, vendor, and/or version of the requesting user agent' (Mazilu, 2022) . In doing this, your requests will become more transparent and trustworthy.

Web-scraping is an increasingly important means of creating/ extracting/ gathering data in the social sciences (Luscombe et al., 2022). The next section will present the code utilised in scraping the International Energy Agency's policy database.

Method

i. Import libraries to build tool.

```
import requests
from bs4 import BeautifulSoup
from tabulate import tabulate
import csv
```

ii. Define Number of pages to scrape.

```
num_pages = 256
```

iii. Create a list of lists to store the policy information.

```
policy_list = []
```

iv. Define the desired keywords (optional).

```
desired_keywords = [minerals]  # Modify or remove this line as needed
```

v. Check if desired_keywords is empty.

```
if not desired_keywords:
    desired_keywords = None
```

vi. Define the User-Agent header (left blank for security purposes)

```
headers = {
    'User-Agent': [redacted]
}
```

vii. Iterate through each page.

```
for page in range(1, num_pages + 1):
```

viii. Send a GET request to the IEA Policies webpage and retrieve the HTML content.

```
    url = f'https://www.iea.org/policies?page={page}'
    response = requests.get(url, headers=headers)
    html_content = response.content
```

ix. Parse the HTML content using BeautifulSoup.

```
    soup = BeautifulSoup(html_content, 'html.parser')
```

x. Find the items under the 'Country' heading.

```
    country_items = soup.select('.m-policy-listing-item__col--country')
```

xi. Find the items under the 'Year' heading.

```
    year_items = soup.select('.m-policy-listing-item__col--year')
```

xii. Find the items under the 'Status' heading.

```
status_items = soup.select('.m-policy-listing-item__col--status')
```

xiii. Find the items under the 'Jurisdiction' heading.

```
jurisdiction_items = soup.select('.m-policy-listing-item__col--jurisdiction')
```

xiv. Find the policy links.

```
policy_links = soup.select('.m-policy-listing-item__link')
```

xv. Iterate over the items under the 'Country' heading.

```
for country, year, status, jurisdiction, link in zip(
    country_items, year_items, status_items, jurisdiction_items, policy_links
):
    country_text = country.get_text(strip=True)
    year_text = year.get_text(strip=True)
    status_text = status.get_text(strip=True)
    jurisdiction_text = jurisdiction.get_text(strip=True)
    policy_title = link.get_text(strip=True)
    policy_url = "https://www.iea.org" + link.get('href')
```

xvi. Check if the item is 'Australia' and (optionally) contains desired keywords.

```
    if (
        country_text == 'Australia'
        and (desired_keywords is None or any(keyword in policy_title.lower() for keyword in desired_keywords))
    ):
```

xvii. Add the policy information to the list.

```
        policy_list.append(
            [country_text, year_text, status_text, jurisdiction_text, policy_title, policy_url]
        )
```

xviii. Print the table of policies corresponding to 'Japan' under the 'Country' heading.

```
if len(policy_list) > 0:
    headers = ['Country', 'Year', 'Status', 'Jurisdiction', 'Policy', 'URL']
    print(tabulate(policy_list, headers=headers, tablefmt='grid'))
else:
    print("No policies found corresponding to 'Japan' from the 'Country' heading.")
```

xix. Save the data to a CSV file.

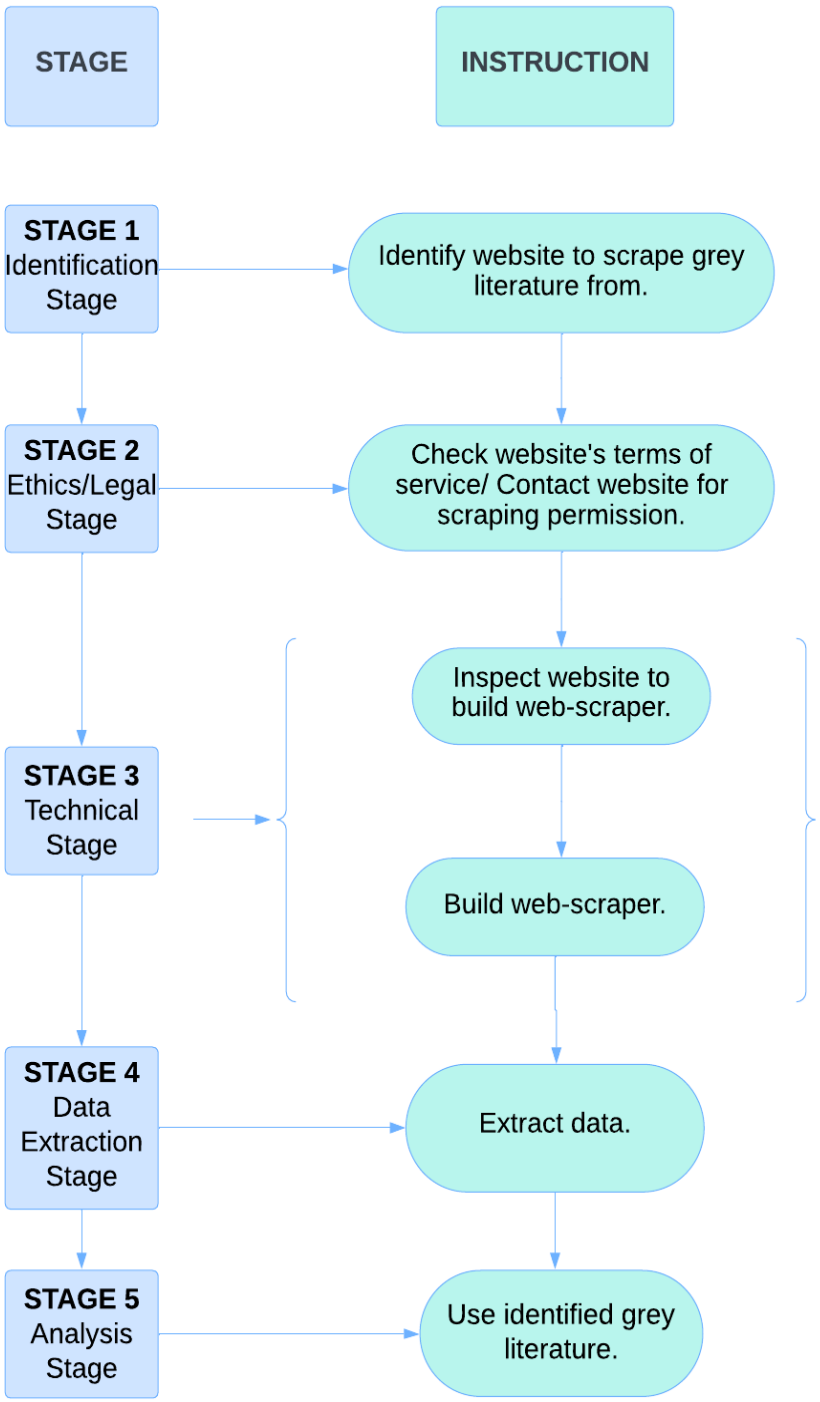
```
csv_file = 'Input file path to save data here'
with open(csv_file, 'w', newline='', encoding='utf-8') as file:
    writer = csv.writer(file)
    writer.writerow(headers)
    writer.writerows(policy_list)
print(f"\nPolicy data saved to {csv_file}.")
```

Results

In creating this method for authors to identify grey literature from a web-based repository, this article identified the different stages that researchers should adhere to as well. These stages are; Stage 1, identifying a suitable webpage for scraping; Stage 2, addressing any ethical or legal concerns before commencing scraping; Stage 3, the technical considerations required to construct a web-scraper for the identified website; Stage 4, Extracting the information (in this case the policy documents); and Stage 5, proceeding to analyse the data obtained through the process. Figure 1

provides an overview of the Identification, Ethical/Legal, Technical, Acquisitional, and Analysis stages of web-scraping in this context.

Figure 1. The different stages for scraping grey literature.



Executing the program led to the following results. As seen in Figures 2 and 3, the program has located all policy documents that contain the word 'minerals' and are from the country 'Australia.' Furthermore, all the information pertaining to the documents, including their URLs have been saved in a CSV file for ease of access, as seen in Figure 4. All data and figures are attributed to © IEA 2023; *Policies database*, www.iea.org/policies, License: CC BY 4.0.

Figure 2. Tabulated data output in Jupiter Notebook (A) .

Country	Year	Status	Jurisdiction	Policy
Australia	2022	In force	National	2022 Critical Minerals Strategy
Australia	2022	In force	National	Critical Minerals Accelerator Initiative
Australia	2022	In force	National	Critical Minerals List
Australia	2022	Announced	International	Minerals Security Partnership
Australia	2022	In force	National	Virtual National Critical Minerals Research and Development Centre
Australia	2021	In force	National	Blockchain Pilot Grants: Critical minerals
Australia	2021	In force	National	Critical Energy Minerals Roadmap
Australia	2021	In force	National	Critical Minerals Facility
Australia	2018	In force	National	Critical Minerals Mapping Initiative
Australia	2001	In force	National	Royalties System on Minerals

Figure 3. Tabulated data output in Jupiter Notebook (B).

URL
https://www.iea.org/policies/15858-2022-critical-minerals-strategy
https://www.iea.org/policies/15863-critical-minerals-accelerator-initiative
https://www.iea.org/policies/15861-critical-minerals-list
https://www.iea.org/policies/16066-minerals-security-partnership
https://www.iea.org/policies/15867-virtual-national-critical-minerals-research-and-development-centre
https://www.iea.org/policies/16651-blockchain-pilot-grants-critical-minerals
https://www.iea.org/policies/15868-critical-energy-minerals-roadmap
https://www.iea.org/policies/15865-critical-minerals-facility
https://www.iea.org/policies/16093-critical-minerals-mapping-initiative
https://www.iea.org/policies/15885-royalties-system-on-minerals

Figure 4. Data as it appears in CSV file.

Country	Year	Status	Jurisdiction	Policy	URL
Australia	2022	In force	National	2022 Critical Minerals Strategy	https://www.iea.org/policies/15858-2022-critical-minerals-strategy
Australia	2022	In force	National	Critical Minerals Accelerator Initiative	https://www.iea.org/policies/15863-critical-minerals-accelerator-initiative
Australia	2022	In force	National	Critical Minerals List	https://www.iea.org/policies/15861-critical-minerals-list
Australia	2022	Announced	International	Minerals Security Partnership	https://www.iea.org/policies/16066-minerals-security-partnership
Australia	2022	In force	National	Virtual National Critical Minerals Research and Development Centre	https://www.iea.org/policies/15867-virtual-national-critical-minerals-research-and-development-centre
Australia	2021	In force	National	Blockchain Pilot Grants: Critical minerals	https://www.iea.org/policies/16651-blockchain-pilot-grants-critical-minerals
Australia	2021	In force	National	Critical Energy Minerals Roadmap	https://www.iea.org/policies/15868-critical-energy-minerals-roadmap
Australia	2021	In force	National	Critical Minerals Facility	https://www.iea.org/policies/15865-critical-minerals-facility
Australia	2018	In force	National	Critical Minerals Mapping Initiative	https://www.iea.org/policies/16093-critical-minerals-mapping-initiative
Australia	2001	In force	National	Royalties System on Minerals	https://www.iea.org/policies/15885-royalties-system-on-minerals

Discussion

Collecting grey literature for a systematic literature review is generally a hard, time consuming, and resource intensive task. However, as this article has illustrated, there are methods available to increase the rigor, speed, transparency, and repeatability of this process. Aimed at researchers that have some coding experience, this article has systematically set out the steps taken to create a web-scraping tool in Python. Unfortunately, formal training in algorithmically-driven techniques like web-scraping is not a high priority in the social sciences (Luscombe et al., 2022). Therefore, undergraduates, and post-graduates alike will need to seek alternative ways to appraise

themselves of these techniques. Fortunately, there is not a steep learning curve in this approach. However, the challenges of web-scraping due to technical, legal, and ethical considerations require constant learning and problem solving (Luscombe et al., 2022). This was evident in the writing of this article. In obtaining permission from the IEA for scraping their database, care had to be taken to ensure that data from partner institutions was not included. As set out in point 4 in the IEAs Terms of Use for the Policies and Measures Databases (PAMS), further permissions would be needed in the event that data from joint databases (such as the IEA/IRENA database) were scraped (IEA, 2023).

There is a process that researchers need to go through to ensure data is obtained both legally and ethically. Nonetheless, by employing web-scraping techniques to scrape large data bases, researchers are presented with opportunities to leverage their research to include grey literature in a rigorous, transparent, and cheap manner.

Conclusion

The open science paradigm implores researchers to develop/utilise methodologies that include the publishing of raw data sets, codes employed, alongside more detailed analyses (Crüwell et al., 2019; Echtler & Häußler, 2018; Shmagun et al., 2022). The outcome of this article, the organisation of policy documents that originate in 'Australia' and contain the keyword 'minerals', could have been compiled manually. Researchers could have utilised the IEA websites' functions to go through and manually select each policy and copy the information into an Excel workbook. However, this is an unstructured and non-systematic way to gather and organise these documents. In the context of the work presented in this article, a method for extracting policy documents from an online repository has been introduced. It's intent is to highlight an approach to make gathering grey literature for inclusion in a systematic review more systematic. When the codes to create the tool is included alongside results, this in turn serves to make gathering grey literature more rigorous, transparent, and importantly, repeatable.

Conflict of interest

The author has none to declare.

Funding Declaration

The author's work is supported by a stipend provided by the University of Tasmania alongside a top-up scholarship provided by Natural Hazards Research Australia. Neither organisation was involved in the preparation, review, or submission of this article.

Ethic Statement.

Ethics was sought from the host institution; however it was not required. All data was obtained with the permission of the IEA. As such all data is © IEA 2023; *Policies database*, www.iea.org/policies, License: CC BY 4.0.

AI Statement.

The author acknowledges the use of CHAT GPT 3.5 in the methods section regarding the creation and refinement of the coding put forth in this article. However, the code has been reviewed by the author, found to be correct, and as such the author stands by the generated results.

References

Adams, J., Hillier-Brown, F. C., Moore, H. J., Lake, A. A., Araujo-Soares, V., White, M., & Summerbell, C. (2016). Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. *Systematic Reviews*, *5*(1), 164. https://doi.org/10.1186/s13643-016-0337-y

Anderljung, M., & Hazell, J. (2023). Protecting Society from AI Misuse: When are Restrictions on Capabilities Warranted? *arXiv preprint arXiv:2303.09377*. https://doi.org/10.48550/arXiv.2303.09377

Astanin, S. (2022). *Tabulate 0.9.0*. Retrieved 24/7/2023 from https://pypi.org/project/tabulate/

Banks, M. (2009). Blog posts and tweets: the next frontier for grey literature. In *Grey literature in library and information studies*. De Gruyter. http://eprints.rclis.org/15411/9/5%2014%20Banks.pdf

Benzies, K. M., Premji, S., Hayden, K. A., & Serrett, K. (2006). State-of-the-evidence reviews: advantages and challenges of including grey literature. *Worldviews on Evidence-Based Nursing*, *3*(2), 55-61. https://doi.org/10.1111/j.1741-6787.2006.00051.x

Berners-Lee, T., Fielding, R., & Frystyk, H. (1996). *Hypertext transfer protocol--HTTP/1.0* (2070-1721). https://www.rfc-editor.org/rfc/rfc1945

Blackhall, K. (2007). Finding studies for inclusion in systematic reviews of interventions for injury prevention–the importance of grey and unpublished literature. *Injury Prevention*, *13*(5), 359-359. https://doi.org/10.1136/ip.2007.017020

Blackwell, A. F. (2002). What is programming? PPIG, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=fffaed8f98da11d4c6fb3d692dc0ffd41736f209

Bogdanski, S., Chang, B.-C., Lawal, I., Baugh, G., Blanton-Kent, B., Coonin, B., Mundle, K., Oddo, A., Riel, S., & Sarcia, R. (2005). Collecting Grey Literature. *Science and Technology Libraries*, *25*(3), 35-70.

Caretta, M. A., Drozdzewski, D., Jokinen, J. C., & Falconer, E. (2018). "Who can play this game?" The lived experiences of doctoral candidates and early career women in the neoliberal university. *Journal of Geography in Higher Education*, *42*(2), 261-275. https://doi.org/10.1080/03098265.2018.1434762

Chandra, R. V., & Varanasi, B. S. (2015). *Python requests essentials*. Packt Publishing Birmingham, UK.

Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J. C., Orben, A., Parsons, S., & Schulte-Mecklenbeck, M. (2019). Seven easy steps to open science. *Zeitschrift für Psychologie*. https://doi.org/10.1027/2151-2604/a000387

Diouf, R., Sarr, E. N., Sall, O., Birregah, B., Bousso, M., & Mbaye, S. N. (2019). Web scraping: state-of-the-art and areas of application. 2019 IEEE International Conference on Big Data (Big Data), http://doi.org/10.1109/BigData47090.2019.9005594

Dongo, I., Cadinale, Y., Aguilera, A., Martínez, F., Quintero, Y., & Barrios, S. (2020). Web scraping versus twitter API: a comparison for a credibility analysis. Proceedings of the 22nd international conference on information integration and web-based applications & services, http://doi.org/10.1145/3428757.3429104

Dreber, A., & Johannesson, M. (2019). Statistical significance and the replication crisis in the social sciences. In *Oxford research encyclopedia of economics and finance*.

Echtler, F., & Häußler, M. (2018). Open source, open science, and the replication crisis in HCI. Extended abstracts of the 2018 CHI conference on human factors in computing systems,

Freelon, D. (2018). Computational Research in the Post-API Age. *Political Communication*, *35*(4), 665-668. https://doi.org/10.1080/10584609.2018.1477506

Glez-Peña, D., Lourenço, A., López-Fernández, H., Reboiro-Jato, M., & Fdez-Riverola, F. (2014). Web scraping technologies in an API world. *Briefings in bioinformatics*, *15*(5), 788-797. https://doi.org/10.1093/bib/bbt026

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

Gunnell, K. E., Belcourt, V. J., Tomasone, J. R., & Weeks, L. C. (2022). Systematic review methods. *International Review of Sport and Exercise Psychology*, *15*(1), 5-29. https://doi.org/10.1080/1750984X.2021.1966823

Gupta, P., & Sehgal, N. K. (2021). *Introduction to machine learning in the cloud with python: Concepts and practices*. Springer Nature.

Haddaway, N. R. (2015). The use of web-scraping software in searching for grey literature. *Grey J*, *11*(3), 186-190. https://www.researchgate.net/profile/Neal-Haddaway/publication/282658358_The_Use_of_Web-scraping_Software_in_Searching_for_Grey_Literature/links/5c3af240458515a4c721fea1/The-Use-of-Web-scraping-Software-in-Searching-for-Grey-Literature.pdf

Haddaway, N. R. (2018). Open Synthesis: on the need for evidence synthesis to embrace Open Science. *Environmental evidence*, *7*(1), 26. https://doi.org/10.1186/s13750-018-0140-4

Haddaway, N. R., Bethel, A., Dicks, L. V., Koricheva, J., Macura, B., Petrokofsky, G., Pullin, A. S., Savilaakso, S., & Stewart, G. B. (2020). Eight problems with literature reviews and how to fix them. *Nature Ecology & Evolution*, *4*(12), 1582-1589. https://doi.org/10.1038/s41559-020-01295-x

Hajba, G. L., & Hajba, G. L. (2018). Using beautiful soup. *Website Scraping With Python: Using BeautifulSoup and Scrapy*, 41-96. https://doi.org/10.1007/978-1-4842-3925-4_3

Hannes, K. (2023). Tactics in Support of the Strategy to Include Artistic Research as Evidence in Systematic Reviews. *International Journal of Qualitative Methods*, *22*, 16094069231176391. https://doi.org/10.1177/16094069231176391

Hensel, P. G. (2021). Reproducibility and replicability crisis: How management compares to psychology and economics–A systematic review of literature. *European Management Journal*, *39*(5), 577-594. https://doi.org/10.1016/j.emj.2021.01.002

IEA, I. E. A. (2023). *Terms of Use for the Policies and Measures Databases (PAMS)*. Online. Retrieved 07/08/2023 from https://www.iea.org/terms/terms-of-use-for-the-policies-and-measures-databases-pams

Kamei, F., Wiese, I., Pinto, G., Ribeiro, M., & Soares, S. (2020). On the use of grey literature: A survey with the brazilian software engineering research community. Proceedings of the XXXIV Brazilian Symposium on Software Engineering, https://doi.org/10.1145/3422392.3422442

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and Software Technology*, *51*(1), 7-15. https://doi.org/10.1016/j.infsof.2008.09.009

Krijnen, D., Bot, R., & Lampropoulos, G. (2014). *Automated Web Scraping APIs*. https://staas.home.xs4all.nl/t/swtr/documents/wt2014_web_scraping.pdf

Krotov, V., Johnson, L., & Silva, L. (2020). Tutorial: Legality and ethics of web scraping. https://doi.org/10.17705/1CAIS.04724

Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A.-M. K., Hammerstrøm, K., & Sathe, N. (2016). Searching for studies: A guide to information retrieval for Campbell. *Campbell Systematic Reviews*. http://www.campbellcollaboration.org/images/Campbell_Methods_Guides_Information_Retrieval.pdf

Lawal, A. M., Sambo, A. S., & BUSARI, A. S. (2023). Inhibiting Factors of Selection and Acquisition of Grey Literature in Universities Libraries. https://digitalcommons.unl.edu/libphilprac/7611/?utm_source=digitalcommons.unl.edu%2Flibphilprac%2F7611&utm_medium=PDF&utm_campaign=PDFCoverPages

Linnenluecke, M. K., Marrone, M., & Singh, A. K. (2020). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management*, *45*(2), 175-194. https://doi.org/10.1177/0312896219877678

Lomborg, S., & Bechmann, A. (2014). Using APIs for data collection on social media. *The Information Society*, *30*(4), 256-265. https://doi.org/10.1080/01972243.2014.915276

Luscombe, A., Dick, K., & Walby, K. (2022). Algorithmic thinking in the public interest: navigating technical, legal, and ethical hurdles to web scraping in the social sciences. *Quality & Quantity*, *56*(3), 1023-1044. https://doi.org/10.1007/s11135-021-01164-0

Mahood, Q., Van Eerd, D., & Irvin, E. (2014). Searching for grey literature for systematic reviews: challenges and benefits. *Research synthesis methods*, *5*(3), 221-234. https://doi.org/10.1002/jrsm.1106

Matthes, E. (2023). *Python crash course: A hands-on, project-based introduction to programming*. no starch press.

Mazilu, M. C. (2022). Web Scraping and Ethics in Automated Data Collection. Education, Research and Business Technologies: Proceedings of 20th International Conference on Informatics in Economy (IE 2021), https://doi.org/10.1007/978-981-16-8866-9_24

Nair, V. G. (2014). *Getting started with beautiful soup*. Packt Publishing Ltd.

Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, *10*(3), 233-240. https://doi.org/10.1111/jebm.12266

Romero, F. (2019). Philosophy of science and the replicability crisis. *Philosophy Compass*, *14*(11), e12633. https://doi.org/10.1111/phc3.12633

Sharif, S. M., Yang, A. Z., Hong, F. W., Yee, C. K., Lin, K. Y., Jailani, A. S., Ali, N. B. Z., & Sararaks, S. (2023). Searching for Grey Literature for Rapid Compilation of Malaysia's Health Systems Response During the Covid-19 Pandemic: Challenges and Lessons Learnt. *Journal of Health and Translational Medicine (JUMMEC)*, *26*(2), 81-88. https://doi.org/10.22452/jummec.vol26no2.9

Sheehan, J. (2023). Open Science and the Transformation of Scholarly Communication. *Grey Journal (TGJ)*, *19*(1). https://research.ebsco.com/c/nprl3q/viewer/pdf/vejnd53tib

Shmagun, H., Shim, J., Choi, K.-N., Shin, S. K., Kim, J., & Oppenheim, C. (2022). Korea's national approach to Open Science: Present and possible future. *Journal of Information Science*, *0*(0), 01655515221107336. https://doi.org/10.1177/01655515221107336

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*, 487-510. https://doi.org/10.1146/annurev-psych-122216-011845

Silva, J. A. T. d., & Nazarovets, S. (2023). Assessment of retracted papers, and their retraction notices, from a cancer journal associated with "paper mills". *Journal of Data and Information Science*, *8*(2), 118-125. https://doi.org/doi:10.2478/jdis-2023-0009

Simova, T. (2023). Preregistration of research for theses–a new standard? *SocArXiv*. https://doi.org/10.31235/osf.io/hmdz4

Stratton, S. J. (2019). Literature reviews: methods and applications. *Prehospital and Disaster Medicine*, *34*(4), 347-349. https://doi.org/10.1017/S1049023X19004588.

The Python Standard Library. (2023). *CSV File Reading and Writing*. Python. Retrieved 24/7/2023 from https://docs.python.org/3/library/csv.html

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, *14*(3), 207-222. https://doi.org/10.1111/1467-8551.00375

Uzun, E., Yerlikaya, T., & Kirat, O. (2018). Comparison of Python libraries used for Web data extraction. *Journal of the Technical University-Sofia Plovdiv Branch, Bulgaria*, *24*, 87-92. https://erdincuzun.com/wp-content/uploads/download/plovdiv_2018_01.pdf

Van Rossum, G., & Drake Jr, F. L. (1995). *Python tutorial* (Vol. 620). Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.

Vasiliev, Y. (2020). *Natural language processing with Python and spaCy: A practical introduction*. No Starch Press.

Willard, W. (2003). *HTML A Beginner's Guids - Second Edition*. McGraw-Hill/Osborne.

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

Twenty-Fifth International Conference on Grey Literature

'Confronting Climate Change with Trusted Grey Resources'

OBA Congres • November 13-14, 2023

Oosterdokskade, Amsterdam, Netherlands

Day One Draft Conference Program

09:00 REGISTRATION DESK OPENS

09:30 OPENING SESSION

Welcome Address

Brian Bales, International Nuclear Information System, INIS-IAEA, Austria

Keynote Address and Opening Address

To be announced

Moderator Day One – *Brian Bales, International Nuclear Information System, INIS-IAEA, Austria*

11:30 SESSION 1 – INCREASING THE VISIBILITY OF TRUSTED GREY RESOURCES

- **Finding, Organizing, Using and Preserving Trusted Grey Literature on Climate Change**
  Toby Green, Coherent Digital, France and Julia Gelfand, University of California, Irvine, United States
- **The Contribution of Grey Literature in Achieving Sustainable Fishing: Case Studies from Cambodia, Indonesia, Philippines and Viet Nam**
  Tamsin Vicary and Maria Kalentsits, Food and Agriculture Organization, FAO of the United Nations, Italy; Daryl L. Superio, Northern Iloilo State University, Philippines
- **The Value of Grey Literature on Climate Change Research in the Philippines: A Bibliometric Analysis of Locally Published Researches**
  Daryl L. Superio, Rizzamila R. Superio, and Rolelyn B. Paredes, Northern Iloilo State University; Erish Estante-Superio, Mary Grace Oliveros, and Joy Geromiano, Southeast Asian Fisheries Development Center; Ariette De Asis, Eileen Estoque, and Edna T. Suganob, Northern Iloilo State University; Martin Floro, Capiz State University; Ethelyn Abaday-Maglangit, Mindanao State University at Naawan, Philippines
- **Realizing the Potential of Grey Literature by Recognizing its Publishers: The PUBGREY Project and the Case of Climate Change**
  Dominic Farace, GreyNet International, Netherlands; Stefania Biagioni and Carlo Carlesi, ISTI-CNR, Italy

14:00 SESSION 2 – DILEMMAS, CHALLENGES, AND ACCURACY OF GREY LITERATURE

- **Collection Development and Maintenance of Accurate Grey Literature on Climate Change: A Case Study of the Law and Policy in the United States**
  Tomas A. Lipinski, University of Wisconsin Milwaukee, United States
- **When Trusted Sources Don't Help Us Address Climate Change: A Grey Dilemma**
  Kathrine A. Henderson, LAC-Group, United States
- **Information, Public Decision-Making, and Climate Change: The Many Roles of Grey Literature**
  Bertrum H. MacDonald et al., Dalhousie University, Canada
- **A Review of French PhD Theses on Sustainable Development**
  Hélène Prost, CNRS – GERiiCO; Joachim Schöpfel, Université de Lille, France

16:00 INTRODUCTION TO CONFERENCE POSTERS AND SPONSOR SHOWCASE

Lightening style presentations of each poster in the main conference hall in advance of the Session on Day Two

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

Twenty-Fifth International Conference on Grey Literature

'Confronting Climate Change with Trusted Grey Resources'

OBA Congres • November 13-14, 2023

Oosterdokskade, Amsterdam, Netherlands

Day Two Draft Conference Program

09:00 REGISTRATION DESK OPENS

09:30 POSTER SESSION AND SPONSOR SHOWCASE

The Poster Session continues on the morning of Day Two, where the presenters meet with delegates and participants in an informal setting. Those presenting conference posters are eligible for the Poster Prize 2023 that will be awarded during the Closing Session. Posters will be judged by a panel of jurors on their innovative content, relevance to the conference, graphic design, and presentation.

The Call for Posters opens on May 1, 2023 and closes on October 15, 2023

Moderator Day Two – *Tamsin Vicary, Food and Agriculture Organization of the United Nations, FAO, Italy*

11:00 SESSION 3 – CLIMATE CHANGE AND GREY LITERATURE: THE INIS-IAEA USE CASE

- **Grey literature on climate change studies at the International Nuclear Information System**
  Viet Phuong Nguyen and Brian Paul Bales, Division of Planning, Information and Knowledge Management, Department of Nuclear Energy, International Atomic Energy Agency, Austria
- **Charcoal Burning in Zambia: User Narratives for Successful and Equitable Information Services**
  Brian Paul Bales, Division of Planning, Information and Knowledge Management, Department of Nuclear Energy, International Atomic Energy Agency, Austria
- **Drops in a bucket: contributions of the IAEA Lise Meitner Library to the INIS database**
  Nicolas Rucks, Nuclear Information Section; Division of Planning, Information and Knowledge Management Department of Nuclear Energy; IAEA Lise Meitner Library, Austria

13:30 PROFESSIONAL DEVELOPMENT FORUM

The Twenty-Fifth International Conference on Grey Literature seeks to bring new content to the annual program. GL25 will provide a platform for initiatives with marked potential for grey literature. A Professional Development Forum will accompany the Plenary and Poster sessions to complete this year's conference program. Information professionals are invited to present and discuss special projects and activities in which their organization is currently engaged. https://textrelease.com/gl25forum.html

In order to participate on the Professional Development Forum, submit a working title and brief abstract to conference@textrelease.com. Placement on the conference program also depends on the time allotted presenters. Upon receipt of your submission, you will be contacted with further details.

15:15 CLOSING SESSION

Report by the Conference Moderators, Presentation of the Poster Prize, and Farewell

16:00 POST CONFERENCE TOUR OF THE AMSTERDAM CENTRAL PUBLIC LIBRARY

TextRelease
GL25 Program and Conference Bureau

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

Twenty-Fifth International Conference on Grey Literature
'Confronting Climate Change with Trusted Grey Resources'

OBA Congres • November 13-14, 2023
Oosterdokskade, Amsterdam, Netherlands

Call for Posters

Title of Poster:	
Author Name(s):	**Phone:**
Organization(s):	**Email:**
Postal Address:	**URL:**
Postal Code – City – Country:	

Guidelines

Persons who seek to present a poster during GL25 are invited to submit an English abstract between 200-250 words. The abstract should describe the project, activity, information product or service. The abstract should likewise include a title, name(s) of the creator(s) and full address information. Abstracts are an important source of information available prior to the conference that is accessible to conference delegates and the international grey literature community.

Due Date for Submission

Timely registration is a guarantee for your placement on the conference program. Abstracts in MSWord should be emailed to conference@textrelease.com on or before **October 15th 2023**. Those submitting poster abstracts will receive verification upon their receipt followed by further guidelines for posters, conference registration, etc.

Poster Presentations

Your **digital** conference poster, its abstract and metadata along with your MP4 pre-recorded poster presentation will be published and openly accessible on the **TIB AV-Portal**. Your poster (.jpg or.pdf) along with the accompanying MP4 video recording (max. 5 min.) should be submitted to conference@textrelease.com no later than **November 1st 2023**.

Your **physical** conference poster (maximum size A0 (841 x 1189 mm / 33.1 x 46.8 in) will be displayed on glass panels in the conference foyer adjoining the registration desk and entrance to the conference hall. Those presenting can bring their poster with them or ship it in advance to the postal address of the conference bureau as shown below.

Poster Prize 2023

Those presenting conference posters are also eligible for the Poster Prize 2023. The winning poster will be announced during the GL25 Closing Session. Posters will be judged by a panel of jurors based on their innovative content, relevance to the conference topics, graphic design, and accompanying abstract. ***Click to view prior winners!***

Poster Prize 2020	Poster Prize 2021	Poster Prize 2022
		

TextRelease
GL25 Program and Conference Bureau
Javastraat 194-HS, 1095 CP Amsterdam, Netherlands
https://textrelease.com • conference@textrelease.com
Tel.+31-20-331.2420

Twenty-Fifth International Conference on Grey Literature
'Confronting Climate Change with Trusted Grey Resources'

OBA Congres • November 13-14, 2023
Oosterdokskade, Amsterdam, Netherlands

Professional Development Forum

Title of Presentation:	
Author Name(s):	**Phone:**
Organization(s):	**Email:**
Postal Address:	**URL:**
Postal Code – City:	**Country:**

Guidelines for Abstracts

Participants presenting during the Professional Development Forum are asked to submit an English language abstract between 300-400 words. The abstract should address their involvement in the field of information and its relatedness and/or potential impact for grey literature. Along with your abstract, please also include a brief biographical note that will be used by way of introduction.

Guidelines for Presentations

Each presenter will have 20 minutes for their presentation, excluding Q&A. In order to guarantee a backup, you are required to submit an **MP4 prerecorded video** of your presentation by **November 5, 2023**. Before the forum begins, present a note to the Day Moderator with at least one pertinent question pertaining to your presentation that can be used during the Q&A and subsequent discussion.

Guidelines for Full-Text Papers

Should the Forum Presenter choose to submit a full-text paper for publication in the conference proceedings, their manuscript should not exceed 25 double-spaced pages. The size of the page can be either A-4 or 8½x11 inches. Credit all source material in your bibliography, references, and/or endnotes. A copy of your full text paper should arrive at the office of TextRelease on or before **December 1, 2023** and a corresponding PDF should accompany the Word file by that same date. Your manuscript will also be eligible for publication in The Grey Journal.

On behalf of the Program Chair and Committee, I look forward to your placement on the Conference Program and remain

With sincere regards,

Dr. Dominic Farace,
Program and Conference Director

TextRelease
GL25 Program and Conference Bureau

Javastraat 194-hs, 1095CP Amterdam, NL
https://textrelease.com
conference@textrelease.com
+31 (0) 20 331.2420

Author Information

Atkinson, Cameron
Cameron Atkinson is a researcher at the University of Tasmania. His research interests include strengthening approaches to evidence-based research and how these approaches can be applied to critical infrastructure policy.
ROR_ID https://ror.org/01nfmeh72
ORCID iD: https://orcid.org/0000-0001-5621-5199
Email: cameron.atkinson@utas.edu.au

Bales, Brian
Brian is currently the Coordinator of the International Nuclear Information System of the IAEA in Vienna. He worked previously for the International Residual Mechanism for Criminal Tribunals in Arusha, Tanzania, as well as a previous posting at the IAEA. He holds a Master's Degree in Information Technology from the University of Denver.
ROR_ID https://ror.org/02zt1gg83
ORCID iD https://orcid.org/0000-0003-1452-2785
Email: b.bales@iaea.org

Biagioni, Stefania
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 (ETRDL) 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
Carlo Carlesi graduated in Computer Science and has worked since 1970 at the IEI (now ISTI) of the Italian National Research Council in Pisa. His interest has been concentrated in many software engineering fields such as: Data Base Systems, Software quality and testing, Administration and UNIX system management, Network security and management, Data Publication, Open Science, Digital Library [Management] Systems and Architectures. He was Head of the "Information Technology Security Service" of the Institute and has participated in many national and international projects, among the most recent: Italian Project for Research in the Antarctica - South Pole Project with the aim to create a Multidisciplinary Integrated Information System. Other of the projects in which he participated include: ERCIM Technical Reference Digital Library - ETRDL Project; PUMA (PUblication MAnagement) a software infrastructure for the management of the institutional/disciplinary archives of the published and grey literature of the CNR Institutes. He is currently a collaborator of the Infrastructures for Science Laboratory (InfraScience) of the CNR ISTI and is involved in the GreyGuide Repository and Portal to Good Practices and Resources in Grey Literature. He has co-authored a number of publications dealing with digital libraries and grey literature. In 2013, he became involved in the GreyGuide Project, where he now provides for the Technical Development and Support. In 2018 together with GreyNet International, DOIs were minted and entered in the metadata fields of GreyGuide records. Carlo's involvement with persistent identifiers carried over to the AccessGrey Project in 2019. He is currently involved in the PID Project that includes the ORCiD and ROR ID together with DOI persistent identifiers in metadata records housed in the GreyGuide Repository.
ROR_ID https://ror.org/05kacka20
ORCID iD https://orcid.org/0000-0001-9808-6268

Curtin, Lisa
Lisa Curtin is the executive manager for the International Council for Scientific & Technical Information (ICSTI). She holds a masters' degree in information science from the University of Tennessee (2020), and a masters' degree in literature from Purdue University (2013).
ROR_ID https://ror.org/01jbdvx51
ORCID ID https://orcid.org/0000-0003-1137-7789
Email: lisa.curtin@icsti.org

Author Information

Farace, Dominic
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

Karim, Daniel
Daniel Karim is PhD Candidate of Science in Management from Bandung Institute of Technology, with a topic of interest in multicultural team and leadership. Professor of Aurik Gustomo is a Professor of People Development at Bandung Institute of Technology. And Anggara Wisesa is an Assistant Professor of People and Knowledge Management Interest Group at Bandung Institute of Technology.

Plank, Margret
Margret Plank is currently the Head of the Competence Centre for Non-Textual Materials at the German National Library of Science and Technology in Hannover (Germany). The aim of the Competence Centre for Non-Textual Materials is to develop emerging tools and services that actively support users in the scientific work process enabling non-textual material such as audiovisual media, 3D objects and research data to be published, found and made available on a permanent basis as easily as textual documents. Previously she was responsible for Information Competence and Usability at the TIB. She has also worked as a researcher at the Institute of Information Studies and Language Technology at the University of Hildesheim. She represents TIB on a number of boards including IFLA Steering Committee Audiovisual and Multimedia Section as well as ICSTI / ITOC. Margret Plank holds a Master degree in information science and media studies from the University of Hildesheim, Germany.
ROR_ID https://ror.org/04aj4c181
ORCID_ID https://orcid.org/0000-0001-8941-7563
Email: margret.plank@tib.eu

Reece, Alistair
Alistair is the Product Manager Platform and Implementations at GeoScienceWorld, a leading content aggregator and journal publisher in the geosciences. Alistair has strategic responsibility for the GSW research platform, in particular ensuring that the platform's functionality meets the needs of the geoscience research community, as well as other stakeholder groups. Alistair was also a Business Analyst and Product Owner for one of the major scholarly content platform companies, with particular interest in XML, requirements gathering, and Agile ways of working.
ROR_ID https://ror.org/00ez32163
ORCID_ID https://orcid.org/0000-0002-2002-5593
Email: reece@geoscienceworld.org

Smith, Plato L.
Plato Smith is the Data Management Librarian at the University of Florida with experience in academic research libraries, digital libraries, and data management. He received his doctorate in the field of Information Science from the School of Information within the College of Communication and Information at Florida State University, Florida's iSchool, Summer 2014. From 2005 to 2012, he was Department Head for the FSU Libraries' Digital Library where he developed, populated, and managed digital collections in the FSU Libraries' digital content management system, DigiNole Repository, and electronic theses and dissertations (ETDs) institutional repository.
ROR_ID https://ror.org/02y3ad647
ORCID_ID https://orcid.org/0000-0003-1814-0151

APO | It's good policy.

45,000+ public policy and research resources

6,400+ Publishing organisations

Our Vision: *Policy and practice is improved through evidence-informed decision-making via a wide range of research, analysis and policy resources.*

Our Purpose: *APO is the 'go to' for policy because we collect, connect and collaborate to make research and analysis accessible and usable.*

MyAPO+

My APO+ offers members a range of features to manage and curate APO's rich repository of policy and research, and to stay up to date with the latest releases.

Adding My APO+ paid membership service is a new step in the evolution of APO. This will allow us to continue to meet the policy community's needs and thrive as a self-sustainable public good. It will also mean that our repository and selection of generalist newsletters (Policy Weekly and Policy Pulse) can remain free.

Subject-based newsletters

Filled with the latest releases and delivered to your inbox every fortnight. See below

Saved searches

Get alerts when a resource meeting your search criteria has been added to APO

Metrics button

Measure your impact by tracking pageviews, downloads and shares of any resource on APO

Read later list

On any search results or resource page, click on the 'Read Later' button

My experts

Keep track of resources by your chosen authors or researchers

Notes for Contributors

Non-Exclusive Rights Agreement

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

Submission Requirements

Manuscripts should not exceed 25 double-spaced typed pages. The size of the page can be either A-4 or 8½x11 inches. Provide the title, author(s) and affiliation(s) followed by your abstract, suggested keywords, and a brief biographical note.

A printout or PDF of the full text of your manuscript should be forwarded to the office of TextRelease. A corresponding MS Word file should either accompany the printed copy or be sent as an attachment by email.

REFERENCE GUIDELINES

General

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

Specific

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

Examples

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

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

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

Review Process

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

Journal Publication and Article Deposit

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

Contents

'Current Developments and Initiatives Involving Grey Literature'