

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

Dominic Farace and **Jerry Frantzen**, GreyNet International, Netherlands
Stefania Biagioni and **Carlo Carlesi**, NeMIS Research Laboratory, ISTI-CNR, Italy
Antonella De Robbio, Eprints in Library and Information Science (E-LIS), Italy
Ana Češarek, University of Ljubljana, Slovenia
Cees de Blaaij, Library of Zeeland, Netherlands
Kiyoshi Ikeda, National Institute of Informatics, Japan
June Crowe, Honorary GreyNet Member and Award Recipient, United States

Abstract

Persistent identifiers such as a DOI¹ for a publication and an ORCID² for an author/researcher can be approached from both the demand-side as well as supply-side of information. It appears however that the former attracts more attention. Here emphasis lies in the access to and preservation of research output. Yet, it is on the supply-side regarding the acquisition of research output that persistent identifiers may by the same token have influence in identifying and populating prospective data archives and repositories. This study will look at the influence persistent identifiers have in securing the acquisition of grey literature for public access.

The goal of this project is twofold. First, to carry out a survey within the grey literature community as to the opinions, uses, and applications of persistent identifiers. And second, to initiate a project geared to populate a new collection housed in the GreyGuide Repository³ by using the DOI as an incentive. Resources in Grey Literature (RGL) is as a generic, multidisciplinary collection that will serve for this purpose.

Using GreyNet's distribution channels and social media, stakeholders in the field of grey literature are invited to enter one or more of their publications in the RGL collection. Each new entry will receive a DOI minted by GreyNet International and further stored and preserved in the DataCite registry⁴. Also, a system generated citation will be added to each new entry in order to facilitate record use. The types of grey literature documents eligible for entry in the RGL collection are numerous⁵.

Brief guidelines for record entry require that it be self-archived using the existing online-template and that both the metadata record and accompanying full-text document(s) are in English. An additional descriptive field does allow for entry in another language. And, a translation of the document can also be uploaded in the repository. Finally, it is understood that by submitting the metadata record and file(s), they become open access compliant under Creative Commons license CC-BY-SA⁶.

The initial phase of the project commenced in April 2019 and closed in October 2019. Records harvested during this period along with the results of the survey will be analyzed in its second phase. In the final phase, the project's outcome will be published. Results should indicate whether the AccessGrey Project be extended to other collections in the GreyGuide, and if this project would be of value to other communities of practice in the field of grey literature.

Introduction

The goal of the project was twofold: 1. To learn the opinions, uses, and applications of persistent identifiers within the grey literature community and 2. to explore the use of persistent identifiers, namely the DOI, in the acquisition of grey literature. The method of approach was first to construct a questionnaire that would be used in an online stakeholder survey among a defined population within GreyNet. And, secondly to initiate a campaign among GreyNet's diverse stakeholders by using the DOI and a system generated citation as incentives to deposit documents in the GreyGuide Repository.

AccessGrey Project – Stakeholder Survey

Survey Questionnaire

A search was first carried out in the Collection of Conference Papers on Grey Literature⁷ using the search terms “persistent identifier” and “DOI”, which retrieved 16 full-text documents. Ten questions were then drafted based on the search results. Nine of the survey questions were standardized and one was open ended. All of the questions however did include a comment field. The questionnaire was then entered in SurveyMonkey⁸ from which a link was generated.

Survey Population

The population of the stakeholder survey was drawn from GreyNet’s Distribution List spanning entries from January 1, 2014 to April 24, 2019 - the date when the link to the online survey became operational. Only personal names with both surname and first name or initial were selected. The total population of survey recipients was 509. During the 5-week period in which the link to the survey was online accessible reminder emails were sent out. In total, there were 56 respondents to the survey accounting for an 11% response rate.

Survey Population	Survey Respondents	Survey Results
509	56	x 11,0%

Survey Results and Shared Analyses

The results of the survey are included as an **Appendix** to this paper. However, to maintain the anonymity of the respondents, responses to Question 10, which include names and email addresses, have been removed. Once the online survey was closed to further response, a data paper was drafted and published alongside the survey data in the DANS Easy Archive. Of the 56 survey respondents, 29 chose to provide their contact details. Those who did, were then invited to analyze the data. Five of those respondents submitted their analysis and rightfully share in the co-authorship of this paper.

Excerpts from three of the five analyses are recorded as follows:

[Excerpt 1]

Persistent identifiers such as DOIs are making research more efficient. Additionally, as the existing protocols become more widely adopted, there will be even more improved access to information. Persistent identifiers are not only useful for identifying data but can also be used to store relationships and point to where other data may be stored. They were developed to prevent link rot and to ensure that objects remain available and unchanged. In this way they improve access to information and increase trust in scholarship and research.

It is assumed that the DOI would be only one factor to be considered as adding quality to a research publication. Other factors such as peer review, citations, impact factor, what other researchers say about the paper, author, date, etc. would also need to be considered to indicate quality of research. The assignation of DOIs to metadata records would not necessarily attract more content providers. However, the Open Access (OA) policies of the repository would definitely play a role in attracting content providers. The Grey literature community of respondents to the survey appear to be on the forefront of knowledge about the importance of using PIDs and DOIs. Research in the field of grey literature and its related data will become increasingly accessible as the research information infrastructure becomes more standardized and widely adopted. (*June Crowe*)

[Excerpt 2]

Researchers often search for references that are listed at the end of relevant articles. Over the years the URL links to some sources such as grey literature - often only accessible online - may not work anymore. If there is a persistent identifier, the access to grey literature remains stable and allows continuous access over time. The process of selecting sources that are scientific and relevant to the work of researchers is becoming more important than the ability to find and collect countless sources. In academic circles, when a mentor or professor checks the references listed and sees the DOIs next to sources that are grey literature, these resources actually “count” (even though in their eyes, grey literature may not be a scientific publication). There is a possibility that persistent identifiers such as DOIs, which are recognized worldwide, would encourage researchers to cite their local sources more often. If this is the case, then in some fields such as education, where work and training are led and informed by government guidelines and evaluations – all of which are grey literature (White et al., 2013) – then the results of this survey are encouraging. DOI by itself does not inform us about the quality of the document or data. It does, however, (in cases like a dissertation and a thesis) increase the possibility of connecting the research data with the thesis, which can be seen as a quality indicator, since the content becomes more scientifically provable when the data are available. The creation of the DOI depends on the policy of the repositories and is not directly connected with the quality of the individual record. If Slovenian repositories added the DOI identifier to each thesis, then all would have a DOI - not only the best ones. However, this is not yet the case. Where we now stand is that less experienced researchers may find it difficult to recognize a trustworthy piece of grey literature. In such cases, an international and well-known identifier such as DOI could assist them. **(Ana Češarek)**

[Excerpt 3]

A DOI is not only a persistent but also actionable, because one can plug it into a web browser and be taken to the identified source. In this way, persistent identifiers are strategic to research data outputs because they can be re-used for new research. The persistent characteristic is a guarantee even if the location of a data file may change when an academic changes institution, or when data archive systems become replaced. Examples, not uncommon to Grey Literature. Concerning the question, whether persistent identifiers serve as an incentive in the acquisition of grey literature, a near 27% of the respondents were uncertain. The comment “Probably right” by one of the respondents may be interpreted as a “selling point” to those who were not certain - given the fact that over 30% of respondents strongly agreed. When asked if a repository or data archive that assigns DOIs to metadata records is more likely to attract content providers – one comment was eloquently formulated “In practice this is the case, but the mere fact of assigning DOI's should not replace the other more intrinsic reasons for content providers to choose a certain repository”. This question is in need of further insights to better understand and decide future choices for repositories and digital platforms. This holds particularly in the open access environment where Grey Literature could be a strong pilot light. When asked in the final survey question to provide contact details along with any other comments or recommendations. It is only after being asked to analyze the results of the survey, do I come to recommend perhaps an online course such as a MOOC (massive open online course) that would deal with the meaning and functions of persistent identifiers, their structure, environments, uses and different types. As but one of the 50+ respondents in the survey, we are all assumed to be interested and somewhat experienced. Imagine all the other authors, librarians, and documentalists who work with grey literature. Training in persistent identifiers such as DOIs and ORCIDs would no doubt prove worthwhile.

(Antonella De Robbio)



AccessGrey Project – Acquisition Campaign

The second part of the project dealt with the acquisition of metadata, full-text records using the DOI and a system generated citation as incentives for authors to deposit their grey literature documents in the GreyGuide Repository.

Acquisition Groundwork

In order to channel records to a multidisciplinary collection in the GreyGuide Repository, the existing online template for the RGL (Resources in Grey Literature) had to be revised to include a DOI metadata field as well as a system generated citation. Since the RGL collection is multidisciplinary, it was decided that records in the earlier GGP (Good Practices in Grey Literature) would be merged with the RGL Collection and DOIs would then be assigned to all existing records.

Acquisition Guidelines

Guidelines reflected in the metadata fields of the online template clearly indicate that submissions rely on self-archiving, that the metadata and full-text are in required in English. While other languages can likewise be included in designated fields. Furthermore, it is understood that all records are open access compliant via the CC-BY-SA License.

Acquisition Strategy

Strategies applied in the acquisition of new records have up until now relied on GreyNet's existing channels, namely its Distribution List, The Who is in Grey Literature⁹, Authors who published in the International Conference Series from 2015-2019, GreyNet's Social Media (Facebook¹⁰ and LinkedIn¹¹), and the International Directory of Organizations in Grey Literature¹².

Acquisition Results

To date, the acquisition of new records has been far less than initially anticipated. The RGL Collection¹³ accounts for only 56 full-text records. However, this sample does indicate that the records are multidisciplinary, they represent works from various sectors of government, academics, business, and NGOs. Furthermore, the sample records are published by some 26 corporate authors, from 13 countries worldwide, and together illustrate 17 different grey literature document types.

AccessGrey Project – Outcome and Way Forward

While the results of the Stakeholder Survey clearly indicate the value of persistent identifiers for grey literature, the campaign for the acquisition of records with the incentive of a DOI and system generated citation has until now been considerably less than expected. Given the amount of technical development that has been invested in the start-up of this project, it is considered worthwhile to extend the duration of the AccessGrey Project into 2020. New strategies for the acquisition of records in the RGL Collection reaching beyond GreyNet's current catchment will need to be considered. And, the GreyGuide Repository in which the

RGL collection is housed should apply for registry in OpenDOAR¹⁴, a quality-assured global directory of academic open access repositories. For it is established that the Open Access (OA) policies of a repository definitely play a role in attracting content providers.

References

¹ <https://www.doi.org/>

² <https://orcid.org/>

³ <http://greyguide.isti.cnr.it/>

⁴ <https://search.datacite.org/works?query=greynet>

⁵ <http://www.greynet.org/greysourceindex/documenttypes.html>

⁶ <https://creativecommons.org/licenses/by-sa/2.0/>

⁷ <http://greyguiderep.isti.cnr.it/listtitoli.php?authority=GreyGuide&collection=GLP&langver=en&RighePag=100>

⁸ <https://www.surveymonkey.com/>

⁹ Who is in Grey Literature,
<http://greyguiderep.isti.cnr.it/listtitoli.php?authority=GreyGuide&collection=BIO&langver=en&RighePag=100>

¹⁰ <https://www.facebook.com/greynetinternational>

¹¹ <https://www.linkedin.com/groups/3718857/>

¹² <http://www.greynet.org/internationaldirectory.html>

¹³ Resources in Grey Literature,
<http://greyguiderep.isti.cnr.it/listtitoli.php?authority=GreyGuide&collection=RGL&langver=en&RighePag=100>

¹⁴ <https://v2.sherpa.ac.uk/opensoar/>

APPENDIX: SURVEY RESULTS

Q1 Persistent identifiers increase access to grey literature

- Answered: 56
- Skipped: 0

Strongly Agree	55.36% (31)
Agree	33.93% (19)
Uncertain	10.71% (6)
Disagree	0.00% (0)
Strongly Disagree	0.00% (0)
TOTAL	56

Comments (5)

They can also distract from the main metadata themselves / Clarify the existence of the material and information itself and guarantee access to unstable content / Only if people know how to use them / Document become more trustworthy for readers / Especially DOIs via CrossRef.

Q2 Persistent identifiers serve as an incentive in the acquisition of grey literature

- Answered: 56
- Skipped: 0

Strongly Agree	30.36% (17)
Agree	39.29% (22)
Uncertain	26.79% (15)
Disagree	3.57% (2)
Strongly Disagree	0.00% (0)
TOTAL	56

Comments (1)

Probably right, may be seen as a "selling point"

Q3 Persistent identifiers increase the citation of grey literature

- Answered: 56
- Skipped: 0

Strongly Agree	55.36% (31)
Agree	33.93% (19)
Uncertain	8.93% (5)
Disagree	0.00% (0)
Strongly Disagree	1.79% (1)
TOTAL	56

Comments (6)

Possibly, but they could also introduce a bias / Although it is one of the methods of increasing citation, improvement of quality is also required at the same time / Perhaps the DOI code / Still a lot of work to be done in encouraging best practice in citation / If we use it (because it is more trustworthy), we have to cite it as well. (logical course) / DOI is helpful for reference (and citation) management.

Q4 Persistent identifiers allow for the preservation of grey literature

- Answered: 56
- Skipped: 0

Strongly Agree	41.07% (23)
Agree	41.07% (23)
Uncertain	14.29% (8)
Disagree	3.57% (2)
Strongly Disagree	0.00% (0)
TOTAL	56

Comments (4)

Digital Preservation is the management and maintenance of digital objects / Allow is maybe the wrong word, perhaps assist? / Preservation depends on the IT team / Probably.

Q5 Persistent identifiers are vital in linking and cross-linking data

- Answered: 56
- Skipped: 0

Strongly Agree	58.93% (33)
Agree	26.79% (15)
Uncertain	12.50% (7)
Disagree	1.79% (1)
Strongly Disagree	0.00% (0)
TOTAL	56

Comments (2)

The term "vital" gives too much emphasis to the phenomenon / DOIs are essential for linking data

Q6 A DOI is a quality indicator that increases the value of grey literature

- Answered: 56
- Skipped: 0

Strongly Agree	30.36% (17)
Agree	33.93% (19)
Uncertain	23.21% (13)
Disagree	10.71% (6)
Strongly Disagree	1.79% (1)
TOTAL	56

Comments (6)

Could be, but I wouldn't vouch for it / It is but it shouldn't be as it isn't really any guarantee of quality / I don't think the DOI is indicative of quality, plenty of peer reviewed content with DOIs gets retracted, so where's the quality aspect there? / For me (I am a young researcher) this is true / For as much as I know it's not a quality indicator / It is an investment, and it adds value via the referencing and linking.

Q7 A repository or data archive that assigns DOIs to metadata records is more likely to attract content providers

- Answered: 56
- Skipped: 0

Strongly Agree	33.93% (19)
Agree	46.43% (26)
Uncertain	12.50% (7)
Disagree	5.36% (3)
Strongly Disagree	1.79% (1)
TOTAL	56

Comments (1)

In practice this is the case, but the mere fact of assigning DOI's should not replace the other more intrinsic reasons for content providers to choose a certain repository.

Q8 Do you have an ORCiD or other author/researcher unique persistent identifier?

- Answered: 56
- Skipped: 0

Yes	66.07% (37)
No	23.21% (13)
Not Applicable	10.71% (6)
TOTAL	56

Comments (5)

<https://researchmap.jp/public/about> / ORCiD / I am going to get it soon / And don't like it / Also other IDs (Web of Science, Scopus, HAL...)

Q9 Does one or more of your publications have an assigned DOI?

- Answered: 54
- Skipped: 2

Yes	74.07% (40)
No	16.67% (9)
– Not Applicable	9.26% (5)
TOTAL	54
Comments (2)	
I am not sure / I don't know	

Q10 Please enter your name, email address, and any other comments or recommendations that would be of benefit to this survey

- Answered: 55
- Skipped: 1

I choose to remain anonymous	47.27% (26)
I include my full name and email address, below	52.73% (29)
TOTAL	55
Comments (29)	