

A grey literature publishing model for the Imperial College London institutional repository

Robyn Price^{a*} and John Murtagh^b

^a*Library Services, Imperial College London, London, UK SW7 2AZ;* ^b*Library Services, Imperial College London, London, UK*

Robyn Price is the Bibliometrics and Indicators Manager at Imperial College London, London, UK. Email: robyn.price@imperial.ac.uk. ORCID: 0000-0001-5776-5256

John Murtagh is the Open Access Manager at Imperial College London, London, UK. ORCID: 0000-0003-3341-3437

This is a preprint of an article whose final and definitive form has been published in *Serials Librarian* (2021) [copyright Robyn Price and John Murtagh]; “A grey literature publishing model for the Imperial College London institutional repository” *Serials Librarian* is available at <https://www.tandfonline.com/doi/full/10.1080/0361526X.2020.1847737>

Abstract

In 2019 we became increasingly aware of authors at Imperial College London choosing to publish grey literature through local website PDF or full text hosting. Recognising the need to improve the institutional open access repository as a venue of choice to publish or co-publish grey literature, we developed a publishing model of identifiers (DOIs and ORCIDs) and metrics (indexing, citations and Altmetric coverage). Some of the incentives already existed in the repository but had not previously been explicitly communicated as benefits; whilst others required technical infrastructure development and scholarly communications education for authors. As of September 2020, a 206% increase in deposit of one type of grey literature has been observed over the previous full year, including Imperial’s influential COVID-19 statistical modelling reports.

Keywords: grey literature; institutional repository; persistent identifiers; bibliometrics; scholarly communications

Introduction

Imperial College London is a research-intensive science, technology and medicine Higher Education institution based in London, England. It produces more than 10,000 research papers annually. Since July 2007, the College has provided an Institutional Repository (IR), named 'Spiral', to archive text-based research outputs and make them available as open access¹. At time of writing, the IR is run on DSpace software (v5.10) and currently holds over 68,000 full-text items. All staff and doctoral students are entitled to deposit any text-based work for which they have authorship and sharing rights in the IR. An institutional mandate passed in 2012 requires research publications, subject to publishers' copyright policies, to be deposited into the IR². From 2016 onwards, a national open access mandate, the REF 2021 open access policy, requires accepted journal articles and conference papers to be deposited in a repository as part of the national research assessment exercise³. With the exception of an institutional policy created in 2012 mandating the deposit of PhD theses; grey literature is exempt from the

¹ Fereshteh Afshari and Richard Jones, "Developing an integrated institutional repository at Imperial College London," *Program: electronic library and information systems* 41, no. 4 (2007): 338-352, doi: 10.1108/00330330710831567.

² "Imperial's open access policy", Imperial College London, accessed September 14, 2020, <https://www.imperial.ac.uk/research-and-innovation/support-for-staff/scholarly-communication/open-access/oa-policy/>.

³ "Guidance on submissions" (REF 2021, 2019), HEFCE accessed September 14, <https://www.ref.ac.uk/publications/guidance-on-submissions-201901/>.

mandate and authors of grey literature at the institution are entirely free to create and share the output through their preferred method.

The IR has always been open to grey literature submissions, but few authors were choosing to publish original grey literature with the IR or even deposit the file in the IR after original publication elsewhere. In 2019 we became aware of multiple grey literature authors publishing by hosting a PDF or full text on local websites owned by their institution or research group. Reasons cited informally to us by some of these authors for not using the IR generally related to the limited visuals and functions of the IR user experience. There are no group branding, no full text option and readers are required to click to download the PDF. In contrast, authors often have more options of style and user function on their own local websites; including option to display the full text on the page, use group branding, control the time of upload (this is important for grey literature sent to the media under a press embargo), flexibility to change the PDF post-publication and easy monitoring of downloads via third party web monitoring services such as Google Analytics. A published survey of academics for reasons not to use an IR for any output type also suggested concerns on copyright, originality, perception of quality, fear of plagiarism and lack of technical ability of the user⁴. We agree with the study's conclusion that low-participation is not necessarily a sign of failure of the IR, rather a sign of a plurality of choice that ultimately empowers academics to choose from a strong range of dissemination venues. However we also

⁴Philip M. Davis and Matthew J. L. Connolly, "Institutional Repositories: Evaluating the

Reasons for Non-use of Cornell University's Installation of DSpace," *D-Lib Magazine* 13, no. 3/4 (2011)

feel that the IR is provided by the institution as a service to authors and if authors are not choosing to use it, the service needs to adapt.

It is also possible that the deposit of grey literature in the IR is negatively affected by the mandates controlling open access deposit of research. The substantial efforts required to ensure compliance of research outputs with the UK's Research Excellence Framework (REF) open access policy 2021 could have left academics and research administrators/librarians under-resourced to fully utilise and promote their IR for grey literature. One recent survey of researchers and IR managers found *“researchers see little value in the Institutional Repository....None of the subjects of this research expressed enthusiasm about the Institutional Repository.”*⁵ The author concludes *“The enforced mandate of the REF has resulted in vastly increased deposit, but this has come at a price – Institutional Repository managers are stretched thin, researchers continue to underappreciate what the repository can do, and institutions undervalue work that does not result in ‘good’ publications”*⁶.

Whilst we strongly value the right of authors to publish through their preferred methods when there is no enforced mandate, we felt that the IR can offer benefits that authors do not receive through their local website publishing method. That method also makes it difficult for the institution to collect accurate data on grey literature outputs and offer research and data stewardship support. Through grey literature, IRs have an opportunity to demonstrate value; rather than replicate commercial journals, they can be

⁵ Carolyn Ten Holter, “The repository, the researcher, and the REF: “It's just compliance, compliance, compliance,” *The Journal of Academic Librarianship* 46, no. 1, (2020): <https://doi.org/10.1016/j.acalib.2019.102079>.

⁶ Ten Holter, “The repository,” 8.

a different type of publisher⁷ which is particularly valuable if the IR is publishing original grey literature that will not be published elsewhere (in journals)⁸. In response to this, a project was conceived to identify areas of existing benefits of the IR and scope new incentives to develop a grey literature publishing model that could result in an increase in volume of grey literature published or co-published by the IR.

A grey literature publishing model

Our concept of an institutional grey literature publishing model borrows from established principles of scholarly communication: perpetual open access, reuse licencing and the use of persistent identifiers (PIDs). It should be noted that like many IRs, the Imperial IR does not offer a peer review process⁹. Typically the grey literature published either do not present large amounts of novel research or findings, or the authors require dissemination faster than a journal or preprint server with a screening process will support. Often the publication PDF can be styled to support the institute's branding and targeted to include non-academic readership such as the public, industry and policymakers. Whilst the grey literature authors are academic staff, they often co-author with institutional Communications Managers or Digital Communications

⁷ Clifford A. Lynch, "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age," *ARL Bimonthly Report* 226: 1-7 (2003)

⁸ Terry M. Owens, "Evolution of a Digital Repository: One Institution's Experience," *Journal of Electronic Resources Librarianship* 23, no. 2 (2011): 142-149, doi: 10.1080/1941126X.2011.576959. p.145

⁹ Clifford A. Lynch, "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age," *ARL Bimonthly Report* 226: 1-7 (2003) p.5

Managers who may be responsible for delivery. The typical publication strategy for these works had been to host PDF files or display full text from the author's group or institution website and sometimes accompanied by a small print run. Our publishing model offers scholarly communication principles and practices as incentives for authors to invest time and resources switching publication method.

Identifiers as an incentive

Digital Object Identifiers (DOIs)

Upon deposit in the IR, all items are assigned a permanent identifier using the Handle System, creating a permanent and interoperable URL that resolves to the item.

However, we sought a second identifier for enhanced coverage with databases and one that we felt authors were more familiar with and associated with scholarly works.

Imperial purchased a set of DOI prefixes as part of the British Library Consortium through the DataCite Registration Agency. The IR was configured to permit direct calls to the DataCite API when an item of grey literature was deposited with satisfactory metadata and file. At the configuration stage a decision about which types of grey literature should be assigned a DOI had to be made. Risk of duplicating existing DOIs and lack of suitability of output type for a DOI were the two factors considered when deciding this.

Working Papers

Although some Working Papers deposited in the IR are original documents, spot checking the files revealed the majority to be Author Accepted Manuscripts or preprints

of items already published in journals. This posed the likelihood that assigning DOIs to Working Papers would duplicate the DOIs created by journal publishers or other repositories/archives for the same documents. This would be in contravention of our commitment to assign DOIs only as unique identifiers. In October 2019 the risk of assigning duplicate DOIs to the existing Working Papers in the IR was calculated to affect a minimum 37.2% of the then current number of deposited Working Papers (percentage of Working Papers deposited with a file (446) and an existing DOI declared by the author on submission (166)). The true figure is likely to be higher reflecting non-declaration of existing DOIs by authors. The decision was therefore taken to exclude the Working Papers output type from the DOI assignment process.

Report

The estimated risk of assigning duplicate DOIs to accepted Reports was calculated to be 18.1% of the then current number of deposited Reports (percentage of Reports deposited with a file (160) and an existing DOI declared by the author on submission (29)). Based on this lower risk we decided to assign DOIs to all new items accepted into the Report category, as well as some existing Reports. We have accepted deposits of items such as Data Reports, Briefing Papers and Discussion Papers under the Report category.

Thesis/Dissertation

Thesis is a controlled entry category for completed Imperial College London PhD theses. They are assigned a DOI as per institutional policy.

DOIs in context

Whilst the configuration of the IR has been successful in assigning DOIs for PhD theses and grey literature; the majority of DOIs assigned by Imperial College are not for grey literature, but for data outputs in a separate digital repository owned by the institution (Research Data Repository)¹⁰ that also has DOI assignment capability. The number of assigned DOIs to grey literature are displayed in Table 1 to give an indication of the scale at which grey literature and other output types in the College’s data and publication repositories are engaging with persistent identifiers.

Acknowledging Lynch’s prediction that “*the ability to make persistent reference to materials in institutional repositories will clearly be critical*”¹¹, we have built identifier infrastructure, as both valuable ends in themselves and means to other ends, such as facilitating some of the metrics described in the Metrics as an incentive section of this paper.

Table 1 DOIs assigned by Imperial College London as of 15.09.2020, accessed at <https://stats.datacite.org/> and reproduced with permission

Prefix	Object Type	Total DOIs assigned
10.14469	Imperial College London SPECTRA digital repository	204,711
10.25560	Imperial College London PhD theses	7,799
10.25561	Imperial College London grey literature	332

¹⁰ Jim Downing, Peter Murray-Rust, Alan P. Tonge, Peter Morgan, Henry S. Rzepa, Fiona Cotterill, Nick Day, and Matt J. Harvey, “SPECTRA: The Deposition and Validation of Primary Chemistry Research Data in Digital Repositories” *Journal of Chemical Information and Modelling* 2008 48 (8), 1571-1581 doi: 10.1021/ci7004737

¹¹ Clifford A. Lynch, “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age,” *ARL Bimonthly Report* 226: 1-7 (2003) p. 6

ORCID

The second type of identifier targeted as an incentive for grey literature deposit was Open Researcher and Contributor ID (ORCID). Imperial was an early adopter and founding member of the ORCID organisation in 2014¹². ORCID provides a publication data source in the institution's Current Research Information System (CRIS). Whilst the institution does not mandate that staff adopt an ORCID as a person identifier; we are aware of at least 3,600 current members of staff, both academic and professional, with an ORCID identified by the CRIS. Grey literature authors can benefit from ORCID in the same way that journal-published research does: improved accuracy and linking between their research outputs and their online research profiles, which can be difficult for grey literature authors publishing outside of journal to database infrastructure. Additionally, the institution can benefit from ORCID adding data on grey literature outputs into our CRIS. This increases the visibility of grey literature in the institution and potentially improves the links between people, outputs and funding¹³. To secure these benefits, authors need to validate an ORCID in the CRIS and set up a 'Read and Write to ORCID' connection from the CRIS to ORCID. This setting is available in our

¹² Reimer, Torsten. 2015. "Your name is not good enough: introducing the ORCID researcher identifier at Imperial College London." *Insights* 28 (3): 76-82.
doi:10.1629/uksg.268

¹³ Haak, Laurel L, Martin Fenner, Laura Paglione, Ed Pentz, and Howard Ratner. "ORCID: a system to uniquely identify researchers." 2012 *Learned Publishing* 25 (4): 259-264. doi:10.1087/20120404.

CRIS software (Symplectic Elements v5.10) and we believe offered by other CRIS software packages. If this connection has been enabled, the item can be automatically sent to the ORCID account to feature in the author's ORCID profile, or in reverse from the ORCID profile to the CRIS database. The local website publishing model cannot integrate with ORCID or the CRIS infrastructure.

As well as exploring adding ORCID fields to metadata and surfaced metadata in the IR, we also began encouraging authors to include their own ORCID data in the About the Authors section of a PDF document. The surfacing of ORCID identifiers in the PDF and linking to ORCID.org in journal articles is best practice amongst journal publishers¹⁴. However, we discovered two practical obstacles trying to implement this at an IR level. The first, that although grey literature is authored by academics, in our experience the staff responsible for the formatting and publication, especially for Reports, are communications or public relations professional staff, who are generally less familiar with ORCID than the academic co-authors. If the person setting the layout template and doing final formatting is not aware of ORCID, it is not likely it will be consistently included. The authors also needed practical support in how to create or confirm ORCIDs, especially if they were adding them on behalf of co-authors; and help in using the ORCID.org branding guidelines to correctly present ORCIDs in print and digital format. This was piloted with one grey literature Report, [<https://spiral.imperial.ac.uk:8443/handle/10044/1/77296>] which presented ORCIDs and ORCID badges in the PDF alongside the author names and biographies. They were also hyperlinked to the author's ORCID URI, using the ORCID brand guidelines, see Figure 1.

¹⁴ Haak, "ORCID" p.261

Sustainable mini-grid systems in refugee camps: A case study of Rwanda

JAVIER BARANDA ALONSO  AND DR PHILIP SANDWELL 

About the authors

Javier Baranda Alonso is a Research Assistant on Humanitarian Energy at the Physics Department at Imperial College London. After completing a MSc in Sustainable Energy Futures at Imperial College London, he has focused his research on the provision of low-carbon energy solutions in developing and humanitarian contexts, particularly interested in the potential of sustainable mini-grids for the provision of improved services and livelihood opportunities.

 <http://orcid.org/0000-0001-8157-6332>

Figure 1 Image of ORCID icon and identifier in grey literature report. Baranda Alonso, Javier, and Sandwell, Philip 'Sustainable mini-grid systems in refugee camps: A case study of Rwanda' © 2020 The authors, produced for The Grantham Institute.

Metrics as an incentive

Whilst the hyper competitive metric culture that pervades journal publishing is fortunately mostly absent from grey literature publishing, it has been observed that due to the absence of metrics “*there are many cases... when organisations are charged with producing reports but have no real way of measuring their impact, including when they are distributed free, do not attract academic citations and their sales cannot be tracked*”¹⁵. Authors of grey literature may wish to understand or demonstrate impact of their works. Whilst others have proposed bespoke methods for calculating grey literature citation impact^{16,17}, the computational nature of these methods are not feasible

¹⁵ David Wilkinson, Pardeep Sud and Mike Thelwall, “Substance without citation: evaluating the online impact of grey literature” *Scientometrics* 98, 797–806 (2014).

<https://doi.org/10.1007/s11192-013-1068-7>

¹⁶ Wilkinson, Sud and Thelwall, “Substance without citation,” 798

¹⁷ Matthew S. Bickley, Kayvan Kousha and Michael Thelwall, “Can the impact of grey literature be assessed? An investigation of UK government publications cited by articles and books” *Scientometrics* (advance online publication) <https://doi.org/10.1007/s11192-020-03628-w>

for the authors in our project accustomed to using services such as Google Analytics for quick access to usage data. We adapted and built easy to access download, citations and alternative metrics into the IR infrastructure.

Indexing and citations

Whilst there is not conclusive evidence of citations as indicators of impact for grey literature in all disciplines and there is no institutional mandate to monitor citations for grey literature, some studies suggest that deposit in an open access repository encourages citations¹⁸¹⁹ and enhanced discovery coverage from other content providers poses another incentive to use the IR²⁰. The IR domain is indexed by Google Scholar, and through this project a review was completed to ensure compliance wherever possible with Google Scholar's guidance on metadata optimisation to support their

¹⁸ Kathleen Shearer, "Promoting Open Knowledge and Open Science Report of the Current State of Repositories" (COAR 2015), <https://www.coar-repositories.org/news-updates/promoting-open-science-and-open-knowledge-current-state-of-repositories/> (accessed September 16, 2020).

¹⁹ Anne Gentil-Beccot, Salvatore Mele, Travis C. Brooks, "Citing and reading behaviours in high-energy physics." *Scientometrics* 345-355 (2010); doi:10.1007/s11192-009-0111-1.

²⁰ Tránsito Ferreras-Fernández, Francisco J. García-Peñalvo, José A. Merlo-Vega. 2015. "Open access repositories as channel of publication scientific grey literature." *TEEM '15: Proceedings of the 3rd International Conference on Technological Ecosystems for Enhancing Multiculturality*. Porto, Portugal: Association for Computing Machinery. 419-426. doi:10.1145/2808580.2808643.

automated searching processes

(<https://scholar.google.com/intl/en/scholar/inclusion.html#indexing>). Consequently, we have increased the coverage of the IR's grey literature from Google Scholar (https://scholar.google.com/scholar?start=0&q=10.25561&hl=en&as_sdt=0,5) to enable much of the IR's deposited outputs to be included in Google Scholar. Consequently, grey literature published through it benefits from the authority of the institutional domain to rank highly in Google Scholar and Google searches. Google Scholar coverage includes citation tracking. Grey literature items that are assigned a DOI can also then be recognised by the Web of Science Cited Reference Search. We understand this service will identify citations to grey literature, even if these items themselves are not indexed by their selective-journal databases if they have a DOI and are cited by a source that is indexed. From this method we identified 164 all-time citations in Web of Science indexed journal to 15 grey literature items published by the Imperial IR. 100% (164 of 164) of the citations to were received in 2020, but only 0.6% (1 of 164) of the citations went to grey literature items published prior to 2020. Of the 15 cited items, 87% (13 out of 15) are COVID-19 grey literature research published 2020.

The IR also provides discoverability through the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (<https://www.openarchives.org/pmh/>), ensuring exposure to other content providers such as CORE, OpenAIRE, RIOXX, BASE-SEARCH (Bielefeld Academic Search Engine), OAISTER (now part of WorldCat), Microsoft Academic, and for Theses to EThOS and DART Europe. The harvesting relationship is described in Figure 2. These are aggregator services that extend reach to potential readers beyond the IR. Other than Google Scholar, grey literature authors in our project were generally unfamiliar with these aggregator services, but communication of the benefits of enhanced aggregation and assurance that

harvesting to these sources is only possible through deposit to the IR can provide further encouragement to use the IR. For authors interested in a download metric, we created a download icon that is a visual display of downloads in the record page of every deposited item in the IR. The download data comes from IRUS, a national aggregation service that provides COUNTER-conformant usage statistics for IR usage in the UK [<https://irus.jisc.ac.uk/>]. The download icon is a widget in the IR record page that displays the IRUS download data, providing an easy to access and trusted download metric akin to that of a journal publisher's COUNTER-conforming statistics²¹.

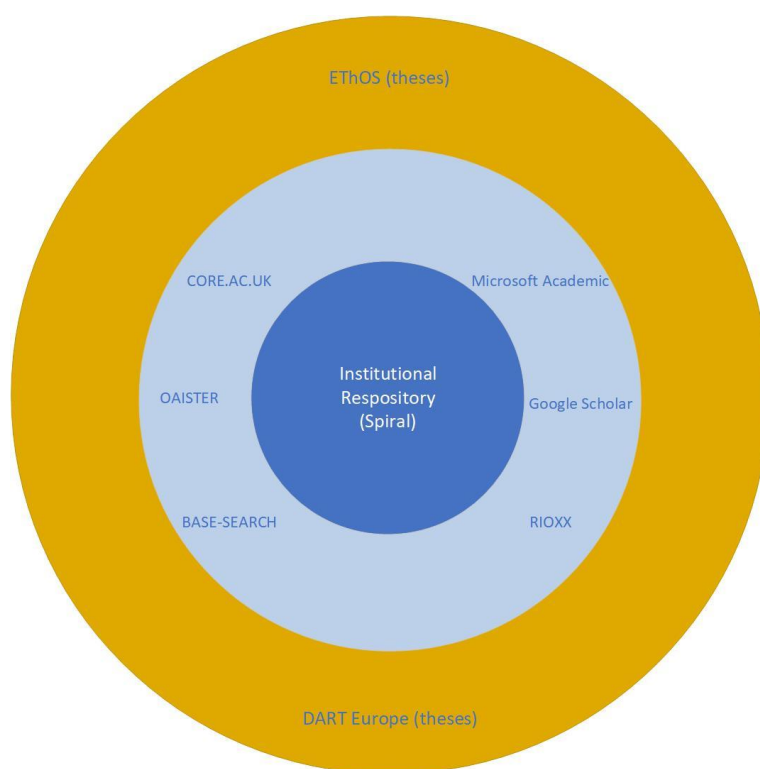


Figure 2 Double layered data harvesting provided by OAI-PMH, diagram based on Figure 2 in Tránsito Ferreras-Fernández, Francisco J. García-Peñalvo, José A. Merlo-Vega "Open access repositories as channel of publication scientific grey literature"

²¹ Ross MacIntyre and Hilary Jones, "IRUS-UK: Improving Understanding of the Value and Impact of Institutional Repositories," *The Serials Librarian* (2016) 70:1-4, 100-105. doi: 10.1080/0361526X.2016.1148423

Alternative metrics

In addition to citations, grey literature authors can use alternative metrics to interpret impact. A well-known commercial alternative metric provider, Altmetric, predominantly tracks journal articles rather than grey literature²², but does have the ability to identify and track grey literature. To identify grey literature items, they must originate from a source domain (such as an IR) that has been proactively added to Altmetric's tracking. It must also contain a persistent identifier, such as a DOI, ISBN, or one of the other types of identifiers listed by Altmetric [https://help.altmetric.com/support/solutions/articles/6000060968-what-outputs-and-sources-does-altmetric-track-]. Publishing via local website hosting cannot fulfil these conditions and any grey literature hosted from these websites would be invisible to Altmetric as an academic source. As Imperial's IR domain was added to Altmetric's list of actively-followed academic sources and DOIs/Handles exists as metadata in the IR page, Altmetric can track mentions to all grey literature if it has been published by the IR.

However, as described in the Introduction, some grey literature authors preferred to co-publish with both the IR and local websites, in order to share their stylised and branded websites on social media and with the press. Generally these local websites display a full text version or link directly to the PDF in one click, skipping the repository publication page. The local-hosting websites do not have the metatags required by Altmetric to identify a scholarly output and track mentions of it. Thus,

²² Cassidy R Sugimoto, Sam Work, Vincent Larivière, and Stefanie Haustein. "Scholarly use of social media and altmetrics: A review of the literature." *Journal of the Association for Information Science and Technology* 68, 9 (2017) 2037-2062. doi:10.1002/asi.23833.

Altmetric is not able to capture activity that came from the local website versions shared online, and this created doubt for authors on the value of using either Altmetric or the IR. After sustained effort from the institution and Altmetric a solution has been derived for authors to publish with IR and share a local website publication page. The solution is writing custom itemprop tags for the item title and DOI into their local website source code. The itemprop tags in effect ‘trick’ Altmetric tracking into aggregating both versions and mentions of both versions into one single and accurate record.

Example of custom metatags inserted into local website hosting an Imperial report:

```
<div itemscope>  
  <meta itemprop="citation_title" content="Climate change and the human-made  
water cycle: Implications for the UK water sector">  
  <meta itemprop="citation_doi" content="10.25561/73992">  
</div>
```

One of the items of grey literature in this project, an Imperial COVID-19 Response group Report co-published by the IR has become Altmetric’s most-mentioned item of all time in their database [https://www.altmetric.com/explorer/outputs?doi_prefix%5B%5D=10.25561&scope=all&show_details=77704842]. For this one item, Altmetric created a custom redirect for their database to capture mentions to the locally-hosted PDF version, thus capturing all the activity. We agree that alternative metrics have been considered to be “widening the

definition of research outputs to include more than just books and journal articles...”²³ and in doing so pose large potential benefit to grey literature authors, by helping bring grey literature to a wider audience²⁴ and broadening interpretation of impact for grey literature.

Supporting grey literature authors

The development of infrastructure for grey literature described has been accompanied by online support resources (<https://www.imperial.ac.uk/research-and-innovation/support-for-staff/scholarly-communication/publishing-with-spiral/>). The resources are aimed at helping grey literature authors adopt scholarly communication best practice. The assumed level of knowledge for this guidance is introductory; accounting for the mix of academic authors who are generally familiar with journal publishing and communications staff responsible for delivery who could be less familiar with these practices. One aspects of grey literature that is generally different to journal published-research, copyright, was identified as a priority for education and support. As per the institution’s Intellectual Property Policy, the institution waives automatic

²³ Euan Adie, “The rise of altmetrics.” In *Altmetrics: a practical guide for librarians, researchers and academics*, ed. Andy Tattersall. (London: Facet Publishing, 2016) 67-82

²⁴ Joachim Schöpfel and Hélène Prost, “Altmetrics and Grey Literature: Perspectives and Challenges.” *GL18 International Conference on Grey Literature*. (New York: 2016)

copyright to research publication and thus items are published by the IR under a default Creative Commons CC BY-NC-ND 4.0 licence, unless the author wishes to apply an alternate. Yet, prior to the project, very few grey literature items were published with any Copyright statement or reuse licensing information in the file. Authors we spoke to were generally unfamiliar with Creative Commons or apprehensive to express rights outside of a traditional journal contributors' agreement, so a lot of grey literature was being published without a licence in the PDF of the document or copyright information in a co-published local website version. We have since provided education resources on interpreting the institutional copyright and applying licences specifically for grey literature [<https://wwwf.imperial.ac.uk/blog/openaccess/2020/08/12/protecting-your-assets-copyright-and-licensing-advice-for-online-reports-briefing-papers-and-working-papers/>] and have observed grey literature published with Creative Commons licensing inserted in the PDF.

The second scholarly communications practice we have adopted for grey literature is citation guidance. Similar to how readers are accustomed to seeing in published journal articles a 'How to cite this document' box with the details of the preferred citation, replicating this has been trialled on some of the grey literature recently published by the IR. Having established DOI-minting for grey literature we would like wherever possible for the DOI to be captured, in the hope that the correct citation will be made by citing authors and picked up by indexers such as Google Scholar and Web of Science Cited Reference Search.

Conclusion

We feel that the essential goal of the project, to make the IR a better option for authors to publish grey literature in, has been accomplished. Developing functionality to assign

DOIs and communicating how to cite the grey literature preceded a rise in the number of Reports deposited to the IR, with 49 deposited between 1st January 2020 and 30th September 2020, a 206.2% increase from 16 in 2019 full year (Figure 3). Reports on the topic of COVID-19 represent 71.4% of the total 2020 deposits (35 out of 49). Many of these authors chose the IR as a co-publication venue rather than the previous method of solely local website publishing. We value this flexibility of choice for authors and the lack of mandate as a unique value of grey literature when so many other types of research dissemination are subject to increasing requirements and compliance demands. Authors can choose to use the IR to benefit from its unique services and don't have to relinquish benefits from other publishing methods if they choose to co-publish. Some of the developments made to the IR may have influenced grey literature authors to choose the IR as their publishing method in 2020, although it would not be possible to ascertain this without surveying authors for their motivations. Monitoring deposit levels in future years will also help determine whether the contributions of COVID-19 research in 2020 to the IR publishing model were an isolated group of deposits or represent long term author preference for the IR. It is not clear whether our project guarantees long term increased deposit of grey literature; but the improved technical infrastructure does offer long term rewards for authors and rewards for the institution in the form of richer data through which to understand grey literature impact. As scholarly technology and IR software evolve there will surely be new ways for IRs to serve authors; current areas of potential development of interest to us include open citations and linking metadata with associated records in other repositories.

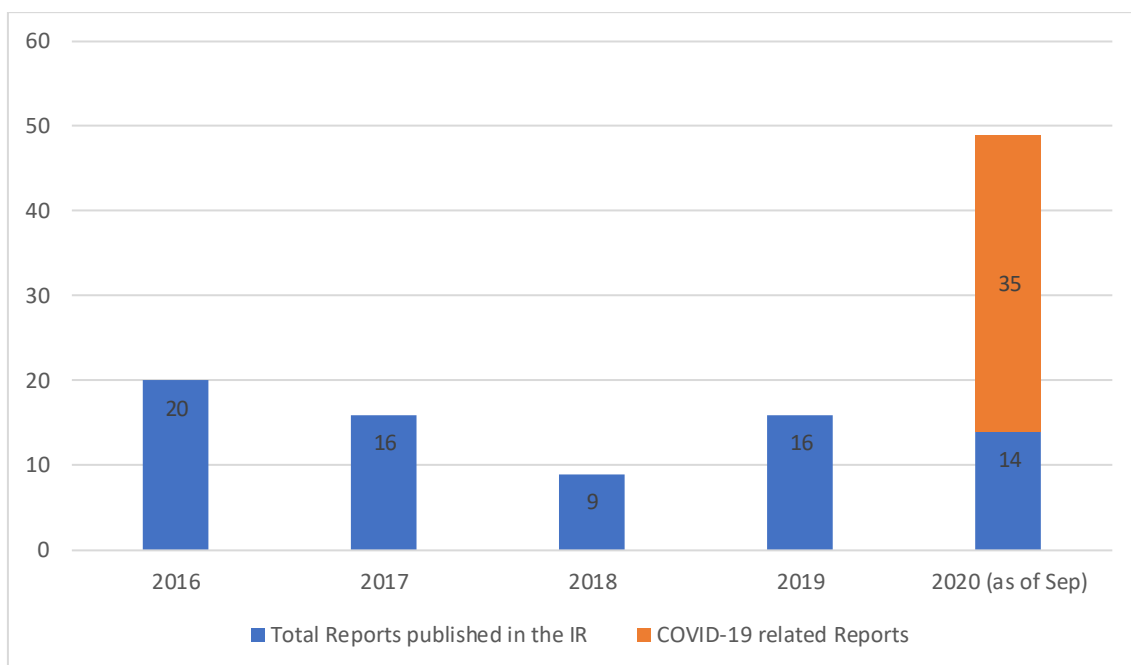


Figure 3 Annual total deposits of Reports in the IR as of September 2020. Reports related to COVID-19 research identified.

Recommendations

- Where there is no mandate, authors are free to publish grey literature through their preferred methods. However, the IR can offer benefits that may not be available through other publishing methods. These benefits need to be developed and communicated to grey literature authors, remembering that not all may be familiar with repositories.
- Search engines, harvesting and aggregator services associated with IRs may increase readership and/or citations for grey literature. Identifiers such as DOIs and support for correct citation practice may result in increased coverage of citations. If assigning DOIs, practicalities of identifying and excluding items that may already have a DOI need to be considered in any workflow.
- Where possible, the IR and CRIS should be integrated with ORCID to permit sharing of grey literature records and linked up data.

- Authors intending to co-publish grey literature between the IR and a local website need to be informed that this could detriment accurate Altmetric coverage. A workaround of custom itemprop metatags can be installed to avert this and permit correct Altmetric tracking of co-published items.
- Where grey literature is produced or co-authored with communications staff, they should be included in scholarly communications education and outreach. A survey of all grey literature authors to understand their needs for research communication and their perspective on the IR could be useful.

Acknowledgements

The authors acknowledge the work of Simon Mackenzie, Andrew McLean and Lalitha Kambhammettu, all of Imperial College London, in developing the IR technical infrastructure described in this project. Emma Proudley and Charlotte Perry-Houts, both of Digital Science, provided feedback on a section of the manuscript and advised the authors on Altmetric and IR integration.

Disclosure Statement

In accordance with Taylor & Francis policy, both authors report that they are employees of Imperial College London. Imperial College London is a licensor of DSpace, Altmetric for Institutions and is a member organisation of ORCID and the British Library DataCite Consortium.

Data Availability Statement

Data supporting this study can be accessed at <http://doi.org/10.5281/zenodo.4075153>.

Bibliography

- Adie, Euan. 2016. "The rise of altmetrics." In *Altmetrics: a practical guide for librarians, researchers and academics*, by Andy Tattersall (Ed.), 67-82. London: Facet Publishing.
- Afshari, Fereshteh, and Richard Jones. 2007. "Developing an integrated institutional repository at Imperial College London." *Program: electronic library and information systems* 41 (4): 338-352. doi:10.1108/00330330710831567.
- Ahcene, Babori, and Aknouche Nabil. 2020. "Increasing the Visibility of Grey Literature in Algerian Institutional Repositories." *Grey Journal* 16 (Special Winter Issue): 43-51. doi:10.17026/dans-z2a-mb5e.
- Bickley, Matthew S., Kayvan Kousha, and Michael Thelwall. 2020. "Can the impact of grey literature be assessed? An investigation of UK government publications cited by articles and books." *Scientometrics*. doi:10.1007/s11192-020-03628-w.
- Davis, Philip M., and Matthew J. L. Connolly. 2007. "Institutional Repositories: Evaluating the reasons for non-use of Cornell University's installation of DSpace." *D-Lib Magazine* 13 (3/4).
- Davis, Philip M., and Matthew J.L. Connolly. 2011. "Institutional Repositories: Evaluating the Reasons for Non-use of Cornell University's Installation of DSpace." *D-Lib Magazine* (3/4).
- Downing, Jim, Peter Murray-Rust, Alan P. Tonge, Peter Morgan, Henry S. Rzepa, Fiona Cotteril, Nick Day, and Matt J. Harvey. 2008. "SPECTRa: The Deposition and Validation of Primary Chemistry Research Data in Digital Repositories." *Journal of Chemical Information and Modeling* 48 (8): 1571-1581. doi:10.1021/ci7004737.

- Fereshteh, Afshari, and Richard Jones. 2007. "Developing an integrated institutional repository at Imperial College London." *Program: electronic library and information systems* 338-352. doi:10.1108/00330330710831567.
- Ferreras-Fernández, Tránsito, Francisco José García-Peñalvo, and Jose A. Merlo Vega. 2015. "Open access repositories as channel of publication scientific grey literature." *TEEM '15: Proceedings of the 3rd International Conference on Technological Ecosystems for Enhancing Multiculturality*. Porto, Portugal: Association for Computing Machinery. 419-426. doi:10.1145/2808580.2808643.
- Gentil-Beccot, Anne, Salvatore Mele, and Travis C. Brooks. 2010. "Citing and reading behaviours in high-energy physics." *Scientometrics* 345-355. doi:https://doi.org/10.1007/s11192-009-0111-1.
- Gerritsma, Wouter. 2015. "Altmetric opportunities for Libraries." *2:AM Altmetrics Conference*. Amsterdam. doi:10.6084/M9.FIGSHARE.1571347.
- Haak, Laurel L, Martin Fenner, Laura Paglione, Ed Pentz, and Howard Ratner. 2012. "ORCID: a system to uniquely identify researchers." *Learned Publishing* 25 (4): 259-264. doi:10.1087/20120404.
- Javier Barando Alonso, Philip Sandwell. 2020. *Sustainable mini-grid systems in refugee camps: A case study of Rwanda*. Report, London: The Grantham Institute, Imperial College London. doi:10.25561/77296.
- Lynch, Clifford A. 2003. "Institutional Repositories: Essential Infrastructure For Scholarship in the Digital Age." *portal: Libraries and the Academy* 3 (2): 327-336. doi:10.1353/pla.2003.0039.
- MacIntyre, Ross, and Hilary Jones. 2016. "IRUS-UK: Improving Understanding of the Value and Impact of Institutional Repositories." *The Serials Librarian* 100-105. doi:10.1080/0361526X.2016.1148423.

- Owens, Terry M. 2011. "Evolution of a Digital Repository: One Institution's Experience." *Journal of Electronic Resources Librarianship* 23 (2): 142-149. doi:10.1080/1941126X.2011.576959.
- Paskin, Norman. 1999. "The digital object identifier system: digital technology meets content management." *Interlending and Document Supply* (MCB University Press) 27 (1): 13-16.
- Reimer, Torsten. 2015. "Your name is not good enough: introducing the ORCID researcher identifier at Imperial College London." *Insights* 28 (3): 76-82. doi:10.1629/uksg.268.
- Schöpfel, Joachim, and Hélène Prost. 2016. "Altmetrics and Grey Literature: Perspectives and Challenges." *GL18 International Conference on Grey Literature*. New York.
- Schöpfel, Joachim, Hélène Prost, and Isabelle Le Bescond. 2011. "Open Is Not Enough: Grey Literature in Institutional Repositories." *GL 13: Thirteenth International Conference on Grey Literature: The Grey Circuit from Social Networking to Wealth Creation. Washington, 5-6 December 2011*.
https://www.researchgate.net/deref/https%3A%2F%2Farchivesic.ccsd.cnrs.fr%2Fsic_00908862.
- Sugimoto, Cassidy R, Sam Work, Vincent Larivière, and Stefanie Haustein. 2017. "Scholarly use of social media and altmetrics: A review of the literature." *Journal of the Association for Information Science and Technology* 68 (9): 2037-2062. doi:10.1002/asi.23833.
- Ten Holter, Caroline. 2020. "The repository, the researcher, and the REF: "It's just compliance, compliance, compliance"." *The Journal of Academic Librarianship* 46 (1). doi:10.1016/j.acalib.2019.102079.

Wilkinson, David, Pardeep Sud, and Mike Thelwall. 2014. "Substance without citation: evaluating the online impact of grey literature." *Scientometrics* 797-806.
doi:10.1007/s11192-013-1068-7.