

# User-Centric Evaluation of Non-Print Legal Deposit in the United Kingdom: The Digital Library Futures Approach

Paul Gooding<sup>1</sup>, Melissa Terras<sup>2</sup>, Linda Berube<sup>3</sup>

(1) University of Glasgow, (2) University of Edinburgh, (3) University of East Anglia

## Introduction

Legal deposit, which ensures the systematic preservation of published materials for future generations, is the legal requirement that a person or group submit copies of official publications to a trusted repository or repositories. Statutory provision for legal deposit dates back to the 16<sup>th</sup> Century (Lariviere, 2000, p. 6), and the concept of legal deposit has existed in English law since 1662, and British law since 1710. The Copyright Act 1911, updated by the Legal Deposit Libraries Act 2003, makes provision for the following six legal deposit libraries to receive copies of publications released in the United Kingdom: the British Library; the National Library of Wales; the National Library of Scotland; the Bodleian Libraries, University of Oxford; the Cambridge University Library; and the Library of Trinity College Dublin. Six years ago, UK legal deposit was extended by “The Legal Deposit Libraries (Non-Print Works) Regulations 2013” (2013) to include online and offline electronic publications in writing, including eBooks, eJournals, electronic mapping, and the UK web domain.

This paper will focus upon usage of Non-Print Legal Deposit (NPLD) collections in UK academic legal deposit libraries. Despite the expansion of UK legal deposit, there is almost no research into the impact and value of NPLD collections for users. Existing studies focus on “four pillars of NPLD strategy: collection development, including selection and metadata; long-term digital preservation of NPLD materials; technical aspects including systems capture, ingest and standards; and regulatory aspects” (Gooding, Terras and Berube, 2019). We therefore aim to address the lack of user-focused evaluation of NPLD services by presenting the work of the AHRC-funded *Digital Library Futures* project. The research set out to answer the following primary research question: what is the impact of the 2013 NPLD regulations upon UK academic deposit libraries and their users? This paper will outline the results of this work, focusing on the challenges that users of NPLD face in order to demonstrate the important role of user-centric evaluation in the development of digital resources. It acts as the first user-centric analysis of the impact and value of NPLD in the UK, and fills an important gap in the methodological literature by demonstrating best practice in studying digital collections for which there is no pre-defined user community.

## Background

Two key points informed our methodological approach. The first was how to account for access arrangements to NPLD. The 2003 Legal Deposit Act, subsequently informed by the 2013 regulations, define how NPLD collections can be accessed and used. The most important points for users are that reader access to NPLD materials is limited only to fixed computer terminals located on the physical premises of the legal deposit libraries, that materials must only be accessible to one user at a time at each legal deposit library, and that forms of reuse such as text and data mining are not allowed. These restrictions are intended to “mirror the level of access to printed publications” (HL Deb, 2013). In practice, there are several areas where this is not the case: for instance, Andrew Green noted that the lack of time limits on the access restrictions created a form of “perpetual copyright” which limits reuse even as materials enter the public domain (2012). Furthermore, the legal deposit regulations specify that any additional use cases are allowable only through primary legislation, which means that there are discrepancies between NPLD protocols and the latest shifts in disabled provision and copyright law. The government’s guidance specifies that access to NPLD materials for visually impaired persons is based upon the Copyright, Designs and Patents Act 1988 as amended by the Copyright (Visually Impaired Persons) Act 2002. In practice, this means that NPLD protocols for disabled peoples now lag behind the 2014 amendments that extended disability regulations to make accessible formats available to all with a recognised disability. Similarly, the 2014 copyright exception to allow non-commercial text and data mining is not reflected in the regulations.

The second key point was how to define impact and value in relation to NPLD. Researchers have engaged in the definition, modelling, and development of methods for studying the impact and value of digital library collections. However, the resultant work often defines impact poorly or not at all (Gooding, Terras and Berube, 2019, p. 15). Similarly, there is a gap in our understanding of evaluation for resources without a defined contemporary user community. The UK government's approach to NPLD focuses upon posterity rather than contemporary usage, while existing models for impact evaluation such as Tanner's Balanced Value Impact Model refer to measuring impact upon an "intended" community (2012, p. 12). The intended community for NPLD is often defined as future researchers, and so there is a need to consider how evaluation of NPLD can contribute to the development of methodological interventions into collections with poorly or undefined user communities.

Impact evaluation in cultural heritage organisations generally distinguishes between intrinsic value (the value something has in and of itself), and instrumental value (the value something has because it helps to achieve or get something). While the former emphasises notions of cultural significance and prestige that are often the focus of debates around legal deposit, instrumentalism ascribes a clear social function to arts and culture. In this paper, we align the instrumental argument closely with the service-driven ethos of contemporary librarianship (e.g. Shera, 1973; Lankes, 2011), by considering how NPLD might allow the library sector to serve existing and future user needs. In light of this approach, we define value and impact for NPLD as follows:

- 1.) **Value** refers to the benefits, or lack thereof, of NPLD collections for libraries and their users.
- 2.) **Impact** refers to the way in which NPLD collections effect change in collection and managing NPLD collections, and in information seeking behaviour (Gooding, Terras and Berube, 2019, p. 16).

## Methodology

We adopted a mixed methods case study approach to analyse the impact of NPLD upon two key stakeholder groups: 1.) academic deposit libraries in the UK and 2.) users of academic deposit libraries in the UK. Data collection occurred between 2017 and 2018, and was undertaken with the support of our project partners: The Bodleian Libraries, University of Oxford; and the Cambridge University Library. We used a range of qualitative and quantitative methods to address our research questions:

- 1.) **Interviews:** we undertook interviews with 36 expert stakeholders, including key figures at academic deposit libraries, academics in related fields, publishing industry representatives, and policymakers. We used semi-structured interviews, with core sets of questions designed for each stakeholder type and library role holder, and further adapted based on research into an individual's skills and experiences. These questions were further mapped to specific research questions, and formed the basis of flexible interviews of roughly one hour per person. The interviews were subsequently qualitatively coded to allow us to evaluate staff and institutional impact. Our approach to coding reflected the three level approach suggested by Hahn (2008): "initial coding", to develop categories and themes; "focused coding" to start grouping categories under one larger category; and "axial coding" to refine ideas with a finer focus.
- 2.) **Surveys:** we surveyed 40 users of the Bodleian Libraries, and 40 users of the Cambridge University Library, recruited through a heterogeneous purposive sample to ensure representation from a broad range of disciplines. The survey was deliberately delivered to a small group of users, as the objective was to gather in-depth feedback that would allow us to link respondents' broader information seeking behaviour to their experiences of NPLD collections. This was partly a response to previous problems identifying and recruiting users of NPLD collections, as users were generally unaware of the difference between NPLD and other digital collections. The survey was designed around a series of tasks that required users to come into contact with NPLD materials, and to report not only their experience of doing so but other aspects of their scholarly information seeking behaviour. This allowed us to analyse how NPLD interacts with broader models for scholarly information seeking and discovery.

- 3.) **Web analytics:** We also undertook two forms of web analytics. First, we undertook web log analysis of usage of NPLD terminals in the academic deposit libraries, which provided headline statistics of usage of materials via NPLD terminals. Second, we undertook a subject-based analysis of datasets of title-level access requests for NPLD materials. This approach was based upon Marcia Bates' observation that scholarly communication practices function differently across domains, and that "these differences do make a difference" (Bates, 1998, p. 1,200). In other words, it should be possible to identify differences in behaviour by studying which subjects are requested by users. We analysed two datasets, spanning the period from 31<sup>st</sup> July 2015 to 31<sup>st</sup> March 2017: metadata for all eBook title requests (91,809 requests at title level); and metadata for all eJournal article requests (36,505 requests at article level). We developed a small Python-based tool, which used the OCLC Classify2 API service to obtain Dewey Decimal (DDC) and Library of Congress (LCC) classmarks for each record. We then discarded unclassified records and analysed the remaining records to identify subject-based patterns of usage of NLD materials.

## Findings

We discovered almost unanimous agreement that NPLD had immense intrinsic value. Interviewees emphasised the prestige and posterity value of NPLD collections and associated legal deposit with the core mission of their libraries. One interviewee described NPLD as a "gold standard", and the capture of online mapping and web archival materials were viewed by several respondents as major steps forward. This agreement between our interviews led us to conclude that NPLD has broad intrinsic value due to its perceived prestige, potential to benefit future researchers, and its role in preserving the UK's published output. It can therefore be considered highly successful at meeting the UK government's objective of preserving non-print materials for future generations (Department for Digital, Culture, Media & Sport, 2019).

However, it was much more difficult to identify the instrumental value of Non-Print Legal Deposit. Library staff were particularly disappointed with the access arrangements, citing the need for researchers to come into the library as damaging to efforts elsewhere to widen participation and usage. Interviewees repeatedly noted two points as particularly damaging: first, that users were forced to access materials in the library reading rooms rather than remotely; and second, that the restricted opportunities for innovative reuse of NPLD materials were a source of frustration for users. Others felt that it was hard to convince readers of the benefits of NPLD when it went against their efforts to "make it easier for people to get access" to purchased collections, while others expressed dissatisfaction with the quality of NPLD copies in comparison to purchased print and digital collections (Gooding, Terras and Berube, 2019, p. 17).

Our interviewees also felt that user requirements for NPLD collections had not been fully considered in the development and implementation of the regulations until a late stage. This was seen to be damaging because of the way that it undermined other efforts to support users. Similarly, it made it difficult to assess the impact of NPLD upon users because we found that there had been no defined success criteria for users. Very little user assessment had been conducted to contextualise access statistics, leaving a gap in how we understand the impact and value of NPLD.

In response, our methods for analysing usage foreground innovative approaches that are designed to foreground the links between primary data collection and the secondary literature. As users were conceptually unfamiliar with NPLD, we focused upon surveying how NPLD fitted into their broader information seeking behaviour rather than engaging users directly in questions about the collections. We established the baseline characteristics of information-seeking behaviour for our sample, and then gave them information seeking tasks that force them to come into contact with NPLD materials for successful completion. While our respondents came from a variety of disciplinary backgrounds, several common features arose:

- 1.) The respondents worked remotely on a personal device, using commercial search engines or library databases to start their search, and using some form of remote authentication to access subscription materials.
- 2.) They were often engaged in work away from the university, including international fieldwork, that necessitated remote access to library resources.
- 3.) Depending on their discipline, they were likely to report working with a set group of electronic resources.
- 4.) They used web archival materials, including the UK Legal Deposit Web Archive, very little or not at all.

- 5.) They sometimes visited central library sites in person, but were more likely to use faculty, department, or subject libraries due to community and relevance.

This profile maps closely to existing studies which show that users increasingly want remote access to resources, at scale, in the formats that they desire, and across a variety of digital resources. In this sense, NPLD access protocols provide a poor match for information seeking behaviours with digital materials. For instance, the NPLD protocols are problematic when mapped against the “scholarly primitives” posed by John Unsworth: “discovering; annotating; comparing; referring; sampling; illustrating; and representing” (2000). NPLD collections actively impede several of these established scholarly workflows.

As a result, when respondents were asked “Would you consider using NPLD materials regularly, 64% of respondents said no. When asked to elaborate, respondents mentioned similar issues to our library staff: remote inaccessibility and the lack of a clear use case for NPLD materials over other resources were cited as two reasons for their decision. That said, for those willing to travel to libraries, NPLD protocols provide support for certain aspects of scholarly research: respondents noted that they preferred to print, save or read discovered resources right away. NPLD allows reading and printing, and it is therefore likely that researchers who overcome their unwillingness to use fixed terminals are reasonably well served in this regard. However, the need for annotation, bibliographic management, and downloading of materials for later reuse and comparison is less well supported, and put our respondents off using NPLD materials. While these findings are informative in themselves, we would argue that it was key to the success of our study to be able to situate responses in relation to existing literature as a way of addressing the lack of defined community from which to derive objectives for evaluation.

Similarly, we found no evidence that NPLD materials were changing the type of resources that were being used. For instance, our subject-level analysis showed that usage of NPLD collections followed long-established disciplinary boundaries. Access requests for NPLD titles in the Arts and Humanities were more common for eBooks than eJournals, while eJournal requests heavily favoured Technology subjects. These observations reflect existing studies that argue that researchers in the Arts and Humanities still see books as a vital source (Stone, 1982, p. 296; Palmer and Cragin, 2008, p. 171), whereas technology and science subjects tend to rely on faster access to new research and thus favour journals (Talja and Maula, 2003).

This process of combining innovative methods for data analysis with deep reference to existing studies allows us to move towards an understanding of how NPLD challenges established user behaviours with digital resources. By comparing user behaviour more broadly to the opportunities for usage of NPLD, we can therefore conclude that the impact of NPLD upon researchers has been limited, due to a combination of access arrangements and limited awareness of NPLD resources. Furthermore, NPLD access is built upon a case that, to a large extent, impedes online information seeking behaviours and scholarly work. We understand such a methodological approach to be user-centric in nature, because it necessitates both close attention to the user context for a particular digital resource, and to the broader scholarly literature surrounding digital resources in general. In doing so, we found that the difficulties associated with evaluating resources in the absence of a defined community were alleviated. That said, it also made clear that a neglect of similar work in NPLD implementation meant that there is no compelling argument that the instrumental value of NPLD has been as fully realised as its intrinsic value.

## **Conclusion**

Non-Print Legal Deposit has been extremely successful in preserving the United Kingdom’s digital textual publications, but it also provides a case study into the implications of neglecting users in the planning and implementation phases of resource development. The case of NPLD indicates how neglect of the user context during strategic planning can lead to digital resources that pose significant challenges for library user communities. We therefore propose that it is necessary to adopt a culture of “User-Centric Evaluation” of NPLD. We propose that user-centric evaluation should be built upon a framework of five key tenets:

- 1.) Users are the long-term beneficiaries of NPLD, not publishers or libraries.
- 2.) The change in information sharing, libraries and research communities that results from the digital turn necessitates re-evaluation of print as a default reference point for NPLD.
- 3.) Publishers are entitled to protect their commercial and legitimate interests, but the growing significance of Open Access and changing Intellectual Property rights cannot be ignored.
- 4.) Libraries must be empowered to take actions in response to emerging information behaviours. These actions should be based on evidenced trends, and focused upon making collections accessible, usable, and meaningful to users in the long term.
- 5.) The first four tenets require continued collaboration between libraries, publishers, and user groups.

These five tenets are expanded upon in our recent white paper (Gooding, Terras and Berube, 2019). However, certain competencies are required to successfully develop such a culture of user-centric evaluation of NPLD. When devising our methodology, we found existing models for impact evaluation to be incomplete because they refer to existing user communities. For NPLD this was reflected in the split between the confident expression of its intrinsic value and the hesitancy over its instrumental value. The problematic status of NPLD for contemporary users was found to stem from the fact that NPLD collections were not seen as a service to contemporary users. The value of NPLD was framed largely in terms of its long-term benefits, but there was no clear plan for ensuring that these benefits were realised in the long term.

In methodological terms, there is a need to address this gap between existing models for impact evaluation, which emphasise impact upon intended communities, and collections for which there is no clearly identifiable community in the short to medium term. Certain activities, such as resource development, interface design, and digital preservation (Digital Preservation Coalition, 2017), require proactive evaluation to ensure they remain relevant and meaningful in the future. We therefore argue that user-centric analysis is key to address this problem, because it aligns evaluation activities with the service-driven ethos of contemporary librarianship. In other words, user-centric analysis where communities are poorly defined requires librarians to consider not only issues relating to technical implementation and UX, but to develop methods that meaningfully relate broader information seeking trends to the resource in question. This can often require a “flexible and potentially experimental approach to research methodology” (Gooding, Terras and Berube, 2019, p. 27) to uncover relationships in innovative ways. It also requires broad engagement between library practitioners and researchers in order to ensure the relevance of theoretical and critical work to practical implementation of library digital resources.

### **Acknowledgements**

The *Digital Library Future* project was funded by the AHRC under grant number AH/P005845 between June 2017 and May 2019.

### **References**

Bates, M. (1998) ‘Indexing and access for digital libraries and the Internet: human, database, and domain factors’, *Journal of the American Society for Information Science*, 49(13).

Department for Digital, Culture, Media & Sport (2019) *Post-Implementation Review of the Legal Deposit Libraries (Non-Print Works) Regulations 2013*. Department for Digital, Culture, Media & Sport. Available at: <https://www.gov.uk/government/publications/post-implementation-review-of-the-legal-deposit-libraries-non-print-works-regulations-2013> (Accessed: 16 July 2019).

Digital Preservation Coalition (2017) *Non-Print Legal Deposit Digital Preservation Review: Final Report*. Digital Preservation Coalition. Available at: <https://www.bl.uk/britishlibrary/~media/bl/global/digital%20preservation/non-print-legal-deposit-digital-preservation-review.pdf> (Accessed: 16 July 2019).

Gooding, P., Terras, M. and Berube, L. (2019) *Towards User-Centric Evaluation of UK Non-Print Legal Deposit: A Digital Library Futures White Paper*. Glasgow: University of Glasgow. Available at: <http://elegaldeposit.org/resources>.

- Green, A. (2012) 'Introducing Electronic Legal Deposit in the UK: A Homeric Tale', *Alexandria: The Journal of National and International Library and Information Issues*, 23(3), pp. 103–109. Available at: <https://journals.sagepub.com/doi/10.7227/ALX.23.3.15> (accessed 16 July 2019)
- Hahn, C. (2008) *Doing Qualitative Research Using Your Computer: A Practical Guide*. London: Sage Publications Ltd.
- HL Deb (2013). Available at: [https://hansard.parliament.uk/Lords/2013-03-06/debates/13030683000250/LegalDepositLibraries\(Non-PrintWorks\)Regulations2013?highlight=%22legal%20deposit%22#contribution-13030683000082](https://hansard.parliament.uk/Lords/2013-03-06/debates/13030683000250/LegalDepositLibraries(Non-PrintWorks)Regulations2013?highlight=%22legal%20deposit%22#contribution-13030683000082) (accessed 16 July 2019).
- Lankes, R. D. (2011) *The Atlas of New Librarianship*. Cambridge Mass.: The MIT Press.
- Lariviere, J. (2000) *Guidelines for Legal Deposit Legislation*. Paris: United Nations Educational, Scientific and Cultural Organization. Available at: <http://unesdoc.unesco.org/images/0012/001214/121413eo.pdf> (Accessed: 16 July 2019).
- Palmer, C. L. and Cragin, M. H. (2008) 'Scholarship and disciplinary practices', *Annual Review of Information Science and Technology*, 42(1), pp. 163–212.
- Shera, J. H. (1973) 'Toward a Theory of Librarianship and Information Science', *Ciência da Informação*, 2(2), pp. 87–97.
- Stone, S. (1982) 'Humanities Scholars: Information Needs and Uses', *Journal of Documentation*, 38(4), pp. 292–313.
- Talja, S. and Maula, H. (2003) 'Reasons for the use and non-use of electronic journals and databases: A domain analytic study in four scholarly disciplines', *Journal of Documentation*, 59(6), pp. 673–691..
- Tanner, S. (2012) *Measuring the Impact of Digital Resources: The Balanced Value Impact Model*. London: King's College London. Available at: [http://www.kdcs.kcl.ac.uk/fileadmin/documents/pubs/BalancedValueImpactModel\\_SimonTanner\\_October2012.pdf](http://www.kdcs.kcl.ac.uk/fileadmin/documents/pubs/BalancedValueImpactModel_SimonTanner_October2012.pdf) (Accessed: 16 July 2019).
- The Legal Deposit Libraries (Non-Print Works) Regulations 2013* (2013). Available at: <http://www.legislation.gov.uk/ukxi/2013/777/contents/made> (Accessed: 16 July 2019).
- Unsworth, J. (2000) 'Scholarly Primitives: What Methods do Humanities Researchers Have in Common, and How Might our Tools Reflect This?', in *Humanities Computing: Formal Methods, Experimental Practice*, King's College London. Available at: <https://www.people.virginia.edu/~jmu2m/Kings.5-00/primitives.html> (Accessed: 16 July 2019).