
The material cannot be used for any other purpose without further permission of the publisher and is for private use only.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

http://eprints.gla.ac.uk/235439/

Deposited on 02 March 2021

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk
Identifying the Future Direction of Legal Deposit in the United Kingdom: The Digital Library Futures Approach

Authors:

Contact: Dr. Paul Gooding (University of Glasgow) paul.gooding@glasgow.ac.uk

Prof. Melissa Terras (University of Edinburgh)

Linda Berube (City University of London)

Publication: Journal of Documentation (Forthcoming)

DOI: https://www.doi.org/10.1108/JD-09-2020-0159

This author accepted manuscript is deposited under a Creative Commons Attribution Non-commercial 4.0 International (CC BY-NC) licence. This means that anyone may distribute, adapt, and build upon the work for non-commercial purposes, subject to full attribution. If you wish to use this manuscript for commercial purposes, please contact permissions@emerald.com.

Acknowledgements

The research that informs this article was supported by a grant from the Arts and Humanities Research Council under grant number AH/P005854 between June 2017 and August 2019. We are grateful to our project partners at the Cambridge University Library, and the Bodleian Libraries, University of Oxford, for their expert support.
Abstract:

- Purpose (mandatory):
  o To date, there has been little research into users of the Legal Deposit Libraries (Non-Print Works) Regulations 2013. This paper addresses that gap by presenting key findings from the AHRC-funded *Digital Library Futures* project. Its purpose is to present a “user-centric” perspective on the potential future impact of the digital collections that are being created under electronic legal deposit regulations.

- Design/methodology/approach (mandatory):
  o The study utilises a mixed methods case study of two academic legal deposit libraries in the United Kingdom: The Bodleian Libraries, University of Oxford; and Cambridge University Library. It combines surveys of users, web log analysis, and expert interviews with librarians and cognate professionals.

- Findings (mandatory):
  o User perspectives on NPLD were not fully considered in the planning and implementation of the 2013 regulations. We present findings from our user survey to show how contemporary tensions between user behaviour and access protocols risk limiting the instrumental value of NPLD collections, which have high perceived legacy value.

- Originality/value (mandatory):
  o This is the first study to address the user context for UK Non-Print Legal Deposit. Its value lies in presenting a research-led user assessment of NPLD, and in proposing “user-centric” analysis as an addition to the existing “four pillars” of legal deposit research.
1.) Introduction

The digital turn has simultaneously transformed access to library collections, and heralded a shift in the production and dissemination of published textual media. This shift now challenges the conventions that underpin legal deposit, a statutory obligation intend to ensure the long-term preservation of a nation’s cultural and intellectual heritage (Larivière, 2000). Legal deposit regulations are widespread, but not ubiquitous: in 2016, De Beer et al. (2016) reported that 62 out of 245 national and state libraries worldwide either benefited from legal deposit regulations or participated in legal deposit activities. Many nations have already extended legal deposit to encompass electronic publications, resulting in large and increasingly significant national digital collections.1 There has been little sign, though, that the rationale behind legal deposit has changed alongside this digital transformation. In this article we will focus upon the United Kingdom, where the aim of legal deposit has remained consistent in recent decades. As the recent Post-Implementation Review of Non-Print Legal Deposit (NPLD) describes, the objective remains to allow for the “preservation of the UK’s non-print publications for future generations” (Department for Digital, Culture, Media & Sport, 2019). Such wording closely replicates that used in parliamentary debates in the 1990s (HC Deb, 1997), and suggests that cultural inheritance has always been a higher priority than the needs of contemporary users. However, many researchers have publicly questioned whether the restrictive access protocols for NPLD are in fact a barrier to the usage of electronic publications.

The AHRC-funded Digital Library Futures (DLF) project, which ran from 2017 to 2019, addressed the lack of research into the impact of NPLD upon the users of the UK Legal Deposit Libraries. The key findings of the project were published in a white paper, where we concluded that “existing models for evaluating collections [are] unsatisfactory for studying NPLD, because they address existing user communities” (Gooding, Terras and Berube, 2019,
This paper therefore aims to address the role of user assessment in tracing the potential impact of electronic legal deposit in the UK and beyond. It will seek to answer the following questions:

- What interventions are required to identify the potential impact of NPLD collections upon future users?
- What might contemporary usage of NPLD collections tell us about the implications of the current access arrangements for future generations?

First, this paper will present a review of the UK NPLD regulations, and the methodological literature, to identify how user assessment might address gaps in our understanding of the impact of NPLD. It will then present findings from the DLF project, to demonstrate how contemporary user behaviour can help to assess the potential future impact of the regulations. Finally, it will draw together these insights to argue that “user-centric evaluation” can address the lack of user assessment in the strategic context for NPLD. In doing so, this paper makes an important intervention into the role of legal deposit after the digital turn by establishing the user as a fifth pillar of legal deposit scholarship alongside regulatory aspects, technical implementation, collections development, and digital preservation.

2.) Literature Review

Legal deposit is a statutory instrument that requires designated content creators to deposit copies of defined publications with a specified national institution (Larivière, 2000, p. 3). Since the introduction of the first formal regulations in the sixteenth century, legal deposit has been widely adopted. It provides a public good that secures long-term access to published works (De Beer et al., 2016, p. 88), and forms the basis of many national library collections (Brazier, 2016, p. 42). ‘Electronic legal deposit’ (‘e-legal deposit’) is the broad term for legal deposit of
digital materials, while ‘Non-Print Legal Deposit’ (NPLD) refers specifically to the UK, because the relevant regulations refer to the deposit of “non-print works” (The Legal Deposit Libraries (Non-Print Works) Regulations 2013). This literature review will outline the development of NPLD in the United Kingdom, before addressing two interrelated bodies of literature. First, it explores the scholarship on NPLD to establish key thematic pillars. Second, it addresses existing practices in the assessment of digital resources, to identify how user assessment is needed to address gaps in our understanding of the long-term user context for NPLD.

2.1) Development of Non-Print Legal Deposit in the United Kingdom

The UK strategy for NPLD arose from the recommendations of the Working Party on Legal Deposit (1998), which was set up in the late 1990s in response to a surge in the publication of born-digital materials that fell outside the purview of the existing legal deposit regulations. The recommendations of the working party laid the framework for voluntary electronic deposit, and for the Legal Deposit Libraries Act 2003 that established the principle of e-legal deposit in UK law. Subsequently, the Legal Deposit Libraries (Non-Print Works) Regulations 2013 enacted NPLD into law, and formalised the arrangements for collecting, preserving, and using deposited non-print materials. The requirement for publishers to deposit non-print materials is additional to the arrangements for printed materials, and covers work “in writing”: any item that is “(a) transmitted by electronic means; (b) received in legible form; and (c) capable of being used for subsequent reference” (The Legal Deposit Libraries (Non-Print Works) Regulations 2013, 2013). The UK legal deposit regulations are not dynamically aligned with related legislation, and so are informed by the Copyright, Designs and Patents Act 1988, without incorporating subsequent amendments.ii
Due to this lack of alignment, the Legal Deposit Libraries Act 2003, informed by the 2013 regulations, defines “what can be collected under NPLD, from whom, and how it can be accessed and used” (Gooding, Terras and Berube, 2019, p. 10). In practice, the access protocols aim to “mirror the level of access to printed publications” (HL Deb, 2013). Reader access to NPLD materials is limited to fixed computer terminals within legal deposit library reading rooms. Materials must be accessible to only one concurrent reader at each legal deposit library. Deposit libraries cannot support text and data mining of NPLD materials, nor make provision for disabled persons other than those with visual impairments (Department for Digital, Culture, Media & Sport, 2019). The regulations therefore provide restrictive access protocols that, it should be noted, are very similar to those in many other nations with advanced e-legal deposit programmes (Alexandrov, 2018). To date, though, there has been little consideration of impact of NPLD upon contemporary or future users.

2.2) The ‘Four Pillars’ of Research into NPLD

The existing literature focuses upon topics that form the “four pillars” (Gooding, Terras and Berube, 2019, p. 6) of research into the impact and implementation of NPLD: regulatory aspects, including observations on the development of the NPLD regulations; aspects of technical implementation, including systems, capture, ingest, and standards; collections development, including selection and metadata; and the long-term digital preservation of NPLD materials. De Beer et al. (2016) have reviewed the challenges that national libraries face in response to e-legal deposit, and many scholars have outlined practices in their own national context (e.g. Larsen, 2005; Mason, 2007; Bródka, 2014; Cadavid, 2014; Degerstedt and Philipson, 2016; Derrot and Koskas, 2016). However, this review will focus primarily upon the United Kingdom, and upon NPLD.
2.21) Regulatory Aspects

Several papers record the development of NPLD in the UK: Geoff Smith (1999) discusses the work of the 1998 Working Group for Legal Deposit; Clive Field (2004) provides an account of the negotiations until the introduction of the Legal Deposit Libraries Act 2003; Richard Gibby and Andrew Green (2008) outline the work of the Legal Deposit Advisory Panel (LDAP) from 2005 to 2008; and Gibby and Caroline Brazier (2012) and Green (2012) provide separate accounts of the path to the finalisation of the 2013 regulations. These papers provide a complete timeline of the workings of the strategic bodies behind NPLD, which explains how the regulations were agreed and introduced. They focus not only on the design of the regulations, but how discussions balanced the “interests” (Field, 2004, p. 98) of publishers and legal deposit libraries: to protect publishers’ intellectual property, and to provide a comprehensive framework for preserving the national digital record, respectively. It is in light of these interests that the conditions for access were defined. Smith notes that publishers’ representatives on the Working Party for Legal Deposit considered that:

Deposit in multiple copies, or networked access at more than one location, would be unacceptably burdensome, could allow control of the data to be lost, and in some cases would endanger the commercial viability of a publication (Smith, 1999, p. 127).

Several studies have been published about the relationship between publishers, libraries, and legal deposit. John Feather’s (1994) historical account of the relationship between publishers and copyright covers the industry’s frequent opposition, while John Davies (1998) provides a useful insight into the tensions between publishers, libraries and universities in the mid-1990s. Both authors make it clear that many publishers, even those supportive of the foundational principles of legal deposit, have found the details of its implementation to be too onerous. Indeed, Adrienne Muir (2020) notes that as far back as the nineteenth century, publisher compliance with the legal deposit act was often unsatisfactory. Muir provides a far-reaching
account of the importance of the publishing trade to the United Kingdom, and its often stormy relationship with legal deposit. She concludes that publishers acknowledge the benefits of legal deposit for preserving the written record, but continue to be concerned that “legal deposit collections could be used in ways that interfere with their commercial interests” (Muir, 2020). Legal deposit enjoys qualified support from the publishing industry, as Byford summarises:

Authors and publishers, especially those involved in academic and professional publishing, share common interests such as long term research value to the nation and their material preserve in a stable and organized environment. The legal deposit libraries are ideally suited to realize those and other objectives (Byford, 2002, p. 294).

Publisher concern about the potential for broader access to legal deposit materials contributes to the decision that NPLD “replicates de jure the de facto physical limitations of reading a printed book” (Gibby and Brazier, 2012, p. 371). Smith (1999) noted that the Working Party publicly recommended that these protections should be permitted for the duration of the copyright period, but this has not been achieved because the NPLD regulations restrict access in perpetuity (Gibby and Brazier, 2012, p. 371). Green (2012, p. 108) describes this as the implementation of “perpetual copyright,” and results in NPLD access protocols that are more restrictive than the equivalent protocols for print materials (Gooding, Terras and Berube, 2019, p. 11).

2.22) Technical Implementation

The introduction of e-legal deposit has contributed to the advancement of technical library infrastructures. Representatives of European national libraries have described the challenges in ensuring metadata quality across e-legal deposit collections (for instance Degerstedt and Philipson, 2016; Derrot and Koskas, 2016). Others have reported on how the UK legal deposit libraries have developed and implemented technical infrastructures for NPLD collections. Milne (2008), for instance, outlines the work that was done by the LDAP to consider the most
suitable technical solutions for heterogeneous materials in the planning phase, including eJournals and the UK web. Others have considered processes for capturing unique content streams, particularly the UK Web Archive (Hockx-Yu, 2014b; Jackson, 2015). These focus upon two aspects: ensuring the scalability of capture and preservation solutions; and ensuring the security and integrity of access solutions. For instance, the UK deposit libraries have published descriptions of ERICOM, the shared system that originally ensured secure delivery of legal deposit materials to library terminals through a “secured remote desktop browser system” (British Library, 2013). The technical solutions are heavily informed by the restrictions of the access protocols, with delivery of the materials focusing primarily upon security and scalability rather than user needs.

2.23) Collection Development

NPLD has had notable effects upon collection development in the legal deposit libraries. Nicholas Joint (2006, pp. 469–472), for instance, describes the post-2003 arrangements for voluntary deposit of non-print publications. He raises issues of scale, quality control and preservation that might influence representativeness and comprehensiveness within a national digital collection. Joint warns that many national libraries will largely preserve the same internationally significant journals; while considering this an acceptable risk due to the geographically distributed nature of e-legal deposit collections, he warns that the richness of each national record may become diluted unless it captures “a culture’s unique national digital output, which will tend to be the anarchic, creative explosion of the national web space in all its mix of glory and goriness” (Joint, 2006, pp. 470–471).

Others have addressed specific subsets of the legal deposit collections. Hannah Chandler (2016) and Jennie Grimshaw (2016), for instance, outline the implications of the regulations for the Official Papers collections at their respective institutions. Both authors see NPLD as an opportunity to address increasingly large gaps in the public record that were caused by previous
restrictions upon scraping web materials. Christopher Fleet and Kimberley C. Kowal (2005) outline the process for capturing and storing digital cartographic data under NPLD. These accounts emphasise that the challenges faced by content specialists are both intellectual and technical in nature, necessitating changes in practice as much as new technical infrastructure. Their emphasis upon data management, software, and hardware demonstrates the complexity of implementing NPLD across a formally heterogeneous national collection comprising eBooks, eJournals, the UK web domain, official publications, maps, and more.

2.24) Long-term Digital Preservation

Finally, the literature focuses upon the long-term preservation of NPLD collections. In 2001, Peter Beagrie (2001) wrote about the challenges that libraries faced as they began to capture digital materials at massive scale. The article, written nearly 20 years ago, predicted the introduction of voluntary and statutory legal deposit to be “very important in securing a framework for the preservation of many of our electronic publications” (Beagrie, 2001, p. 224). In the intervening period, digital preservation has become a notably systemic and collaborative challenge, with the legal deposit community able to share expertise and good practice with the global community. Accounts of this international picture have emerged to support the development of good practice (Mason, 2007) and record the implications of the fast-changing copyright and deposit context of legal deposit collections (Besek et al., 2008). In the UK, there exists a robust picture of best practice in the preservation of NPLD materials. Day et al. (2014) outline the process undertaken to profile collections and define a digital preservation strategy for NPLD. In 2017, the Digital Preservation Coalition (DPC) undertook an independent review of the digital preservation actions undertaken by the legal deposit libraries to deliver their commitment to NPLD. The report concluded that much of the digital preservation practice was “exemplary,” noting that the expertise of the British Library had been “wisely and consistently brought to bear on the preservation of NPLD collections” (Digital Preservation Coalition, 2017,
p. 2). However, the report warned that a lack of user feedback due to restricted user was a risk to quality assurance of digital preservation actions (Digital Preservation Coalition, 2017, p. 23).

2.3) Methods for Evaluating the Impact of Non-Print Legal Deposit

The DPC review supports the notion that the otherwise robust NPLD implementation process has occurred largely absent of close attention to the needs of users. Strategic discussions apparently emphasised a balance between preservation and intellectual property rights, while viewing users largely in the abstract. The UK situation reflects a patchy global understanding of users of legal deposit collections. Just two published studies have addressed access to collections created via e-legal deposit: Helen Hockx-Yu (2014b) explores scholarly use of the UK Web Archive; while Georgi Alexandrov (2018) investigates public access regulations across several European national libraries. Alexandrov explains “e-reading” to describe access protocols for e-legal deposit materials, a process of online search via a fixed library terminal followed by on-screen reading, and explores how European nations with advanced e-legal deposit have achieved this use case. Hockx-Yu highlights a “misalignment between legal requirements and user expectations” (2014a, p. 114) as a challenge for web archives, because of difficulties in making users understand why something as seemingly public as archived websites cannot be accessed online. However, neither study undertakes primary research into the users of e-legal deposit materials, and Hockx-Yu refers to the period before the implementation of the 2013 regulations. There is therefore almost no empirical public research into the users of contemporary e-legal deposit, which the lack of user assessment in the planning and implementation phases of NPLD in the United Kingdom. This lack of attention to users is symptomatic of the status of user research in the heritage sector. The problem is longstanding, with very little in-depth user research that draws evidence from a wide range of sources, and a lack of established metrics across the sector (OCLC Research, 2015).
Previous studies have tried to address this challenge. Meyer *et al.* (2009), for instance, produced the now-retired Toolkit for the Impact of Digitised Scholarly Resources, a comprehensive suite of methods and tools for measuring the impact of digital collections. The resource is now offline, and only available via the Wayback Machine.iii Simon Tanner’s (2012) Balanced Value Impact Model provides a conceptual framework for undertaking impact evaluation in relation to digital resources. Both these models can be seen as a response to the instrumental turn in measuring the value of the cultural heritage sector, which began in the 1980s and was accelerated by New Labour’s cultural shift towards metricisation (Crossick and Kaszynska, 2016, pp. 16–19). As such, both focus explicitly upon supporting institutions to plan for, track, and analyse strategically defined outcomes as evidence to decision makers (Tanner, 2012, p. 38). This is problematic for resources such as NPLD, which have not been the subject of robust strategic planning processes relating to usage. For legal deposit, publishers and policymakers clearly see the beneficiaries of legal deposit as ‘future users,’ and so potential problems inherent in the current regulations can be dismissed as challenges for the future. For instance, in a 2013 House of Lords debate, Lord Gardiner of Kimble responded to criticisms of the “perpetual copyright” clauses as follows:

This is an important issue, but will only arise once the copyright term has ceased, so in practice the issue will not affect legal deposit for many years to come (HL Deb, 2013).

Responses such as Lord Gardiner’s imply merely that benefit accrues from the existence of certain collections, without necessarily considering their future reception. In this way, contemporary values might become embedded in mechanisms for preservation that stop important future processes of remediating, redefining and reinterpreting cultural heritage.iv

Existing models for impact explain future impact via concepts such as “inheritance value” or “historical value” (Throsby, 2003); in other words, deriving satisfaction from the perceived
benefits of bequests to future generations. Paul Knights considers intergenerational bequests in relation to natural landscapes, but his argument is relevant to cultural heritage:

Our posterity-related desires concerning this legacy include the desires that our successors value, preserve and, where necessary, restore the objects that we leave them, continue the projects, traditions and institutions that we pass on to them; and endeavour to understand and appreciate our values, deeds and characters (Knights, 2014, p. 102).

Such interventions frame the aspirations for future beneficiaries in terms aligned with the strategic priorities of contemporary organisations. For instance, the recent Europeana Impact Playbook proposes a model to embed impact assessment into planning and delivery of projects. This positive step encourages organisations to think of impact assessment as an essential feedback loop in planning for impact. The Playbook refers to impact as “changes that occur for stakeholders or in society as a result of activities (for which the organization is accountable)” (Verwayen et al., 2017), a definition based on Tanner’s (2012) Balanced Value Impact Model. However, those impacts are largely defined by internal institutional indicators, and such models are thus not clearly suited to exploring benefits where there is little evidence of an initial strategic context, as is the case here.

In summary, very little work addresses the impact of NPLD upon the users of legal deposit libraries, either globally or in the United Kingdom, despite a broad cross-sectoral agreement of the long-term value of legal deposit and the need to serve smaller research communities. Instead of evaluating NPLD through existing institution-centric paradigms designed to measure impact, which may replicate the existing focus upon the existing ‘four pillars’ of NPLD strategy, the focus should instead turn towards understanding how the regulations affect users and usage, and how this might help to establish usage as a fifth strategic pillar for NPLD scholarship. The article therefore explores the methodological decisions taken by the DLF project, in order to demonstrate how current trends in usage of NPLD collections
might tell us about the problems inherent in the regulations; and how a user-centric evaluative approach to NPLD might be taken forward constructively by the stakeholder community.

3.) Developing a User-Centric Methodology for Non-Print Legal Deposit

The Digital Library Futures project sought to understand how the posterity-driven priorities of NPLD interact with contemporary usage patterns. In light of the research context, two things informed our approach to assessing the impact of NPLD collections: first, how to define impact and value in relation to NPLD; and second, how to analyse user behaviour as a predictor of future impact. Both points relate to the lack of clear and defined contemporary user community for NPLD, or strategic consideration of those users. Instead, we had to consider how existing approaches to user analysis could be leveraged to derive meaningful insights for future NPLD impact assessment and strategy. We used a case-study, Reflection-in-Action approach allowing us to identify “features of the practice situation – complexity, uncertainty, instability, uniqueness and value conflict” (Schön, 1983, p. 18), and an Action Research recursive methodology (Stringer, 2013) in order to understand the complex nature of this previously undocumented area, synthesising data from a variety of sources including interviews, surveys, and data-led user analytics, which we discuss below.

We first sought to define impact and value in relation to NPLD. Researchers have engaged in defining, modelling, and developing methods for studying the impact and value of digital library collections. However, concepts such as legacy are defined in terms of how society currently values or prioritises its own heritage, not in the potential effects of policy decisions for future researchers. However, inherited collections impose a duty of care on present and future generations that requires organisations to consider what exactly is being
inherited. It is already clear that contemporary decisions relating to access, interface design, metadata, and digital preservation actions, embed certain patterns of usage in the short to medium term (Whitelaw, 2015) that will potentially reverberate into the future. User assessment for NPLD must therefore consider the future impacts of contemporary decisions, through the lens of contemporary usage.

The second key point is how to develop methodological insights that successfully address the long-term effects of NPLD. During preliminary discussions, library staff reported that they found it difficult to identify patrons who reported as “users” of NPLD collections. This caused two problems: first, there was no clear user community to form the basis of local assessment activities; and second, there was no existing user benchmarking available to inform external research. However, the restrictive access arrangements mean that contemporary trends in usage of NPLD can be seen as predictive of future use. The legal stasis in which NPLD is held allows us to predict that patterns of usage are likely to be less changeable over time that with less restrictive digital resources. The decisions we took were therefore largely inspired by the definition of impact assessment provided by the International Association for Impact Assessment, as the “process of identifying the future consequences of a current or proposed action” (2009).

Undertaking user assessment in line with this existing definition offers the chance to understand how the regulations affect users of NPLD resources, and to define good practice for digital resources by defining the desired end point for the legal deposit regulations and assessing how regulatory actions might contribute to, or act against, reaching that desired point. That end is a question for negotiation by stakeholders in government, publishing, libraries, the scholarly community, and the public. Our first step towards this was to explicitly frame our definitions of value and impact in forward-looking terms. Cultural heritage impact evaluation 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, instrumentalism ascribes a clear social function to arts and culture (O’Brien, 2010). To consider the future impact of NPLD, though, stakeholders must consider what it allows users to do that they otherwise could not, and how this intersects with observable changes in research and user behaviour. In light of this, we defined value and impact for NPLD as follows:

1.) **Value**: the benefits, or lack thereof, of NPLD collections for libraries and their users;

2.) **Impact**: the way in which NPLD collections effect change in collecting and managing legal deposit collections, and in users’ information seeking behaviours (Gooding, Terras and Berube, 2019, p. 16).

We drew upon a robust suite of existing methods to situate usage of NPLD collections in relation to broader scholarly trends. 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 case study partners at the Bodleian Libraries, University of Oxford, and the Cambridge University Library. We combined robust and well-established qualitative methods with exploratory data science approaches, synthesising the results via Reflection-in-Action (Schön 1983):

1.) **Interviews**: we undertook 36 expert stakeholder interviews, including representatives academic deposit libraries, academic research, publishing, and policymakers. We used semi-structured interviews with core sets of questions for different types of roles, which were then adapted based on each individual’s skills and experience. These questions were mapped to specific research questions to form the basis of interviews of roughly one hour per person. The resultant interviews were transcribed and coded to evaluate
staff and institutional impact, and to assess the status of internal user evaluation for NPLD. Our qualitative coding utilised Hahn’s (2008) three-level model, with two researchers coding independently and then co-developing a final in-depth coding structure.

2.) **Surveys:** we surveyed 40 users of the Bodleian Libraries, and 40 users of the Cambridge University Library, recruited by heterogeneous purposive sampling to ensure representation from across the academic disciplines of each university. The survey was delivered to a small group of users, with the objective of gathering in-depth feedback that would situate NPLD usage within broader behaviours. We designed a series of survey tasks that led users into contact with NPLD materials, reporting not only on that experience but on how it intersected with their information seeking behaviours and preferences. These in-depth findings formed the basis of actionable insights into users of NPLD.

3.) **Web analytics:** we 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 for reading room usage. Second, we undertook a subject-based analysis of title-level access requests for NPLD materials, informed by Marcia Bates’ (1998, p.1200) observation that scholarly communication practices function differently across domains. In other words, it should be possible to identify which scholarly domains might be accessing NPLD collections by examining the frequency with which certain subjects were accessed. We developed a Python-based tool that queried the OCLC Classify2 API service to obtain Dewey Decimal and Library of Congress
classmarks for each record, where possible. After discarding unclassified records, we undertook data analysis to identify subjects used within NPLD collections.¹

While our data collection was built upon using evidence of contemporary usage to predict future usage, it should be noted that such predictions assume the regulations remain largely unchanged. Changes to NPLD regulations and implementation can be reasonably expected to cause changes. This is where it becomes necessary to embed user-centric evaluation alongside other forms of impact assessment as the basis for longitudinal engagement with the policies and practices surrounding NPLD. The following section will elaborate upon what our findings show about the direction of travel for NPLD in its current form.

4.) Users of NPLD: The Neglected Fifth Pillar

Our interviews demonstrated the positive impact of NPLD upon the legal deposit libraries, insofar as it has enabled them to preserve digital materials in unprecedented breadth and depth, and in increasingly diverse formats. Staff interviewees largely felt that legal deposit was a key component in delivering the strategic objectives of their institutions, and that NPLD regulations were a positive addition, as one interviewee noted: “I think it’s one of the highest level strategic priorities for the library. It’s not just about the future of our collections, but also the future of some of our intellectual production in the UK.” NPLD, according to another interviewee, is “seen as a gold standard… by certainly a lot of national libraries, particularly ones that don’t have it.” Several staff interviewees singled out the UK Legal Deposit Web Archive (UKLDWA) as evidence of a collection that was previously impossible for the libraries to capture at scale. They were also conscious of the national mission of legal deposit, using

¹ The development of this tool, and results of this study were expanded upon in a conference paper in 2018 (Gooding et al., 2019), but it will be referenced here insofar as it is demonstrative of the highly exploratory approach that we refer to in this article.
expressions like “public good” and “preserving it for the nation” and thereby echoing the key priorities stated by the UK government. Staff focused upon the benefits of systematic collecting, the diversity of resultant digital collections, and the perceived inheritance value of NPLD collections, as one interviewee noted: “I think it’s not so much now, but in 50, 100 years, 200 years’ time. They’ll be looking at this corpus of material and think it’s absolutely fantastic.”

We found evidence that publishers, librarians and other stakeholders felt NPLD collections possessed a broad legacy value derived from the benefits for future researchers. In this regard, NPLD clearly meets the government objective of capturing digital publications for future generations. However, library staff expressed disappointment with NPLD access protocols. They noted two weaknesses: first, that access was tied to reading rooms; and second, that the protocols contradicted their efforts to widen access to collections, thereby frustrating readers. Their specific criticisms included unintuitive interfaces, items that were inferior in quality compared to subscription and print resources, pagination problems, and difficulties in convincing readers to overlook these perceived downsides (Gooding, Terras and Berube, 2019, pp. 17–18). Our interviewees also felt that very little effort had been made to establish strategic priorities for usage of NPLD. Very little user assessment had been conducted to contextualise access statistics, and internal studies had instead focused upon User Experience (UX) with the NPLD interface. One interviewee noted that users were barely considered in the implementation phase:

When we first started thinking about implementation, we were so consumed with the collection of the material… at no point is the end user considered. The end user isn’t considered in the regulations so it’s never been at the forefront. And all of a sudden we start implementing and realise that the user isn’t represented at all.
Furthermore, several library staff reported that usage of NPLD collections seemed lower than they would expect. The libraries argued similarly in the annexe to the 2019 Post-Implementation Review, noting that the UKLDWA received a fraction of the traffic to the Open UK Web Archive (UKWA) despite being several magnitudes larger. Interviewee anecdotes suggested lower than expected usage, including one who noted that capacity for concurrent users of NPLD resources had not been reached:

At the moment, we handle up to 50 concurrent users. And we probably, maybe get half of that during peak times. So we still have a bit of capacity, and there is scope to scale out a solution to allow a lot more capacity, but we just haven’t had the need to at this stage.

Another interviewee reported that they had carefully observed usage of NPLD terminals in their library, and found no need to increase provision:

We started out with two [terminals] and it was a long time before we saw these being used, and then when we saw two being used at once so we added a third, and then maybe one time we saw all three in use and so we bumped it up to six. I don’t know that I have seen more than three people using them at a time and often there will be nobody there at all.

Of course, this anecdotal reportage exists absent of any recorded expectations for usage. As a strategic decision, this absence is damaging because we are unable to ascertain what was intended for NPLD. In the absence of a clear user strategy, we broached the issue by using our survey as a space to engage with how users of each library felt about NPLD, and how it fitted within their existing information seeking practices. To this end, we invited 40 users from each case study partner to participate. Because there were very few members of the user community who identify as “users” of NPLD, we sought to place their response to these collections in a broader framework of their information needs and behaviour. Our respondents were drawn from a diverse disciplinary background, including STEM, Arts and Humanities, Social
Sciences, and the Medical Sciences. Despite these differences, we identified several commonalities to users’ information seeking behaviour:

1.) Works remotely on a personal device, using commercial search engines or library databases to start their search, using some form of authentication to access subscription materials.

2.) Often engaged in work away from the university, including international fieldwork.

3.) Depending on discipline, is likely to work with a set group of electronic resources.

4.) Uses web archives, including the UKLDWA, very little or not at all.

5.) Sometimes visits central library sites in person, but more likely to use faculty, department or subject libraries due to community and relevance. (Gooding, Terras and Berube, 2019).

Overall, 70% of respondents said that they were not aware of NPLD, with similar awareness levels at each institution: 29 out of 40 at the Cambridge University Library, and 27 out of 40 at the Bodleian Libraries. There was also a bias towards using non-library resources for discovery. When asked which tool they used first for scholarly information seeking, respondents ranked library resources third after browsers and specialist search engines:
Figure 1: Table showing where survey respondents generally start their scholarly information seeking task.

While this is important because these non-library resources will shape user expectations (Nielsen, 2000), it also shows how NPLD fails to fit into existing information practices. Library resources are the only way to search for NPLD materials, whereas most respondents demonstrated a clear preference for other search engines. Furthermore, whereas NPLD is accessible only at fixed terminals on-site, most respondents reported using laptop and mobile devices, and generally did so outside the library. Taking their experiences into account, we asked users whether they would consider using NPLD materials regularly. The vast majority of respondents either said no, or gave a qualified response:
When asked to elaborate, respondents cited inaccessibility and the lack of a clear need for NPLD over other resources as two main reasons for their reticence. Several also noted that their view would necessarily change if the NPLD copy was the only one available to them. One respondent expanded upon how the inconvenience of reading room access bothered them:

If there were a small number of materials for which remote access wasn’t available, then I would accept that I needed to be based in the library to use those resources. However, if it were a regular occurrence then I would find it annoying to not have remote access, because typically I do not use the resources online for a large amount of time, and it would be annoying to make the journey to the library for around 10-20 minutes of use.

In the main, users were clear that remote access was critical to their research. A majority indicated that they valued remote access so highly that they would not bother with a title at all.
unless they considered it critical to the success of their research. This reticence has echoes of Zipf’s “principle of least effort,” which states that “each individual will adopt a course of action that will involve the expenditure of the probably least average of his work” (1949, p. 561). Zipf suggests that users undertake an informal cost-benefit analysis, a constant trade-off between the effort required to employ a particular information seeking strategy and the expect quality or effect of the resulting action. Case and Given (2016), too, note that information seekers will often attempt to minimise the effort required to obtain information, even if it means accepting a lower quantity or quality of information. This should not be mistaken for laziness, however, as our respondents noted they would change their mind if NPLD was the only option; a conclusion that echoes the LAIRAH team’s finding that humanities researcher will spend a lot more time than expected engaging with resources if they believe the results to be beneficial to their research (Warwick et al., 2008).

Finally, user preferences clearly played a role in respondents view of NPLD collections. The following chart demonstrates the factors that users considered when deciding which electronic resources to use:
The most common factors were either not possible with NPLD collections or mirrored accessibility issues which the library interviewees had identified. 76 out of 80 respondents considered remote access in their decision, 61 considered usability, 59 considered the ability to download data, and 51 considered the type of access that is allowed. That said, aspects of the NPLD access protocols did coincide with what users wanted to do. When asked what they would do with resources that they discovered, the overwhelming majority of respondents noted that they either read them right away, saved them, or printed them to read later. While saving
resources is an issue, reading and printing are both provided for NPLD collections at all libraries.

Overall, users were somewhat ambivalent about the benefits of NPLD to them. They were put off by the access protocols, and found the systems relatively difficult to use. The reticence to use NPLD collections, unless materials were otherwise available, furthermore suggests that the impact of NPLD upon researchers has been somewhat limited. The access protocols position NPLD as a resource of last resort, by creating a negative perception of NPLD materials, and by impeding users’ preferred information seeking behaviours. Some stakeholders may see this as the correct approach, but we would like to return to our initial question: what might the experience of our respondents tell us about the implications of the regulations for future users? We can say with some confidence that awareness of NPLD is low, and that users are currently unenthusiastic about the restrictive access protocols. Without changes to the regulations, the gap between user expectations, and what is possible with NPLD collections, is likely to grow with time. Our study supports the idea that users primarily start their scholarly information seeking outside the library ecosystem, and that this is driven by the need for remote access. As we write, these observations are given added urgency by the ongoing Covid-19 crisis. It is therefore reasonable to assume that material which is undiscoverable outside library catalogues, and inaccessible outside library reading rooms, will become increasingly irrelevant to researchers unless it is their only option. If the access protocols are not adapted in time, they will continue to impede future generations of researchers with two likely effects: first, that the intergenerational investment into NPLD infrastructure will deliver an instrumental value that is less fully realised than its perceived legacy value; and second, that significant numbers of users will only engage with NPLD materials when it either becomes the most efficient way of accessing materials despite these barriers, or when their need for that material outweighs the negatives of using the NPLD platform.
It should be noted, though, that users are still coming to terms with some of the implications of the new forms of the digital age. This is evident with web archival materials, which are under-utilised by researchers more generally. While we found that our respondents used eBook and eJournal materials in their research, 65 out of 80 respondents did not use web archives at all. Therefore, while access protocols are certainly one cause of the low uptake of the UKLDWA, low levels of interest also contribute. As a result, NPLD usage must be assessed in a way that would track two factors: first, the potential impact of possible changes to the legislation; and second, the implications of changing user expectations, knowledge and needs over time. It is this careful balancing of the implications of NPLD with broader contextual factors that makes the need for user-centric analysis so pressing.

There are significant challenges that must be addressed to achieve this. First, there is little evidence of user analysis in the development and implementation of the NPLD regulations, and precious little data available on the users of NPLD collections. Moving forward, it is essential that users are considered in strategic discussions, and that a data collection policy is put in place to facilitate longitudinal analysis and decision making. We think of this in terms of establishing a “fifth pillar” of scholarship on legal deposit, centred upon the use and users of e-legal deposit collections. One limitation of any study of this nature is that there may be biases in terms of what is assessed, and by whom. Given that we found a distinct lack of survey respondents who were aware of the concept of NPLD, we believe that the first step towards addressing this is to incorporate user voices into the strategic planning phase of NPLD in future. In line with the action research approach taken, we have begun active talks with the legal deposit libraries regarding the establishment of a user forum to aid this process. The development of user forums to amplify the user voice in NPLD planning is essential, as is increased transparency about usage of NPLD collections across the legal deposit libraries. In terms of concrete actions relating to methodology, we also recognise that in
addition to the limitations of impact assessment, the existing user assessment techniques we use here could usefully be augmented by more experimental approaches that either foreground new insights into existing data, or engage with varied stakeholders to define new approaches to the question of NPLD usage. We therefore propose three future directions for research:

- The development of models for impact assessment of NPLD that ground themselves in exploratory analysis of future impact.
- The foundation of a network on NPLD usage, which will support collaborative exploration of new methods to analyse impact, and train researchers and library professionals in their necessity and application.
- The application of research methods such as data analytics, future-thinking methodologies, and scenario analysis to support a more exploratory and open approach to impact analysis of NPLD.

Finally, the success of any action to improve our understanding of the users of NPLD requires a willingness among stakeholders to engage with the questions of posterity that are seen to be essential to its purpose. This should take the shape of open dialogue between the Legal Deposit Libraries, publishers, and scholars, and these groups should look to establish an ongoing programme of user assessment that can support decision making at a strategic level.

5.) Conclusion: Users as the “Fifth Pillar” of NPLD scholarship

This paper has reported on how the DLF project has approached the challenge of evaluating the potential impact of NPLD collections for future users. By presenting the results of a survey of users of NPLD collections, it demonstrates that significant barriers to the adoption of NPLD exist in the United Kingdom. In doing so, we aim to establish user analysis as the “fifth pillar”
of legal deposit strategy alongside regulatory aspects, technical implementation, collections development, and digital preservation. We have argued here that there is a growing gap between user expectations, and the firmly print-centric ethos of the UK legal deposit access protocols. The lack of alignment with other relevant legislation allows us to predict that these tensions will increase with time, as user needs become increasingly informed by online information seeking behaviour. However, while our user assessment of NPLD collections has provided a clear account of the challenges posed by the existing regulatory environment, the route to change in response is one of advocacy, and of conflicting values. We therefore hope that, rather than delivering a definitive account of the impact and use of NPLD, this study will act as a baseline for a significant increase in activity in user assessment research. This would go a long way towards addressing the current imbalance we noted in the literature around legal deposit.

As we have previously argued, a user-centric approach to NPLD recognises users as long-term beneficiaries, embraces the diverse forms of digital media, entitles publishers to protect legitimate commercial interests, empowers libraries to respond to emerging user needs, and demands continued collaboration between stakeholders in NPLD (Gooding, Terras and Berube, 2019, p. 30). These foundational principles are relevant to the need for deeper methodological engagement with impact assessment as a predictive activity concerned with the potential impacts of current activities (International Association for Impact Assessment, 2009). This paper aims to shape a more user-literate, and digitally literate, approach to legal deposit policy in the future. However, we are conscious that tensions exist between the needs of publishers, libraries and users (Department for Digital, Culture, Media & Sport, 2019). Our response to this is that collaborative, longitudinal analysis of users has the potential to add transparency to these discussions, without privileging an immediate liberalisation of the access protocols. Two areas for future research stand out, however. First, we recognise that more
needs to be done to refine methodological approaches to these challenges. We have not addressed here, for instance, how marginalised potential user communities might be engaged in measuring from the community into the institutional and strategic contexts of NPLD. While we have provided advocacy-based proposals, a clearer agenda around the contribution of various communities needs to be developed collaboratively across relevant keyholders. Second, more work needs to be done to establish a sustainable pipeline of longitudinal impact assessment of NPLD, incorporating diverse perspectives in a wider conversation about the purpose and objectives of the regulations in the digital age.

NPLD is currently a conservative response to the emergence of digital publication. The entire community must go beyond its current position of merely valuing legal deposit as a public good, and towards modelling future usage, in order to break down the misguided assumption that print materials are a default format on which to base policies of access and reuse. The debate must expand beyond issues of contemporary usage, to embrace the question of what society wants legal deposit to be for in the long term. We believe that collaborative action is necessary to achieve this goal, and that it should be informed by careful assessment of the impact of legal deposit regulations. In order to fully realise the benefits of NPLD collections, user-centric research must be embedded at the heart of national policy. Such analysis should interrogate the impact of access protocols on emerging use cases, draw upon existing knowledge of wider shifts in user behaviour, and involve significant collaboration between researchers, publishers, publics, and libraries to promote and understand the unique context within which NPLD collections exist. The future of legal deposit, then, is not an abstract challenge to be addressed by its future beneficiaries, but a methodological and intellectual challenge for the here and now.
References


The importance of these processes has been emphasised in 2020 as the heritage sector responds to the Black Lives Matter movement, which is calling upon institutions and collections to re-evaluate the representation and importance of race in heritage collections, with libraries called upon to decolonise their holdings (Pagowsky and Wallace, 2015; American Library Association, 2020; Charr, 2020).

---

1 Linda Arnold-Stratford and Richard Ovenden (2020) provide a rigorous account of the cope and size of the UK Non-Print Legal Deposit collections in their forthcoming chapter. The collection includes 12Tb of geo-visual data, and 56,000 born-digital music scores, and has seen over 900 publishers transition to a process of depositing electronic materials via the British Library portal.

2 For a fuller account of the development of Non-Print Legal Deposit in the United Kingdom, we recommend the more detailed insider account provided by Richard Gibby and Caroline Brazier (Gibby and Brazier, 2012).

3 The archived website can be found at the following link: https://web.archive.org/web/20091020170038/http://microsites.oii.ox.ac.uk/tidsr/.

4 The importance of these processes has been emphasised in 2020 as the heritage sector responds to the Black Lives Matter movement, which is calling upon institutions and collections to re-evaluate the representation and importance of race in heritage collections, with libraries called upon to decolonise their holdings (Pagowsky and Wallace, 2015; American Library Association, 2020; Charr, 2020).