The Influence of Legal Deposit Legislation on the Digital Collections of the National Library of Scotland

Paul Cunnea, Gill Hamilton, Graeme Hawley, Fred Saunderson (National Library of Scotland)

This is a preprint of a chapter accepted for publication by Facet Publishing. This extract has been taken from the authors' original manuscript and has not been edited. The definitive version of this piece may be found in 'Electronic Legal Deposit, Shaping the Library Collections of the Future,' Facet, London (ISBN 9781783303779), which can be purchased from http://www.facetpublishing.co.uk/title.php?id=303779 The author agrees not to update the preprint or replace it with the published version of the chapter.

Our titles have wide appeal across the UK and internationally and we are keen to see our authors content translated into foreign languages and welcome requests from publishers. World rights for translation are available for many of our titles. To date our books have been translated into over 25 languages.

The National Library of Scotland, with its centres in Edinburgh and Glasgow, is one of six legal deposit libraries in the United Kingdom and Ireland tasked with responsibility for collecting British published print and non-print resources under the Legal Deposit Libraries Act 2003 ¹and the Legal Deposit Libraries (Non-print Works) Regulations 2013². In this chapter we give an overview of the history of the National Library of Scotland, and the growing importance and scale of its digital collections, while contextualising the challenges to access and format that come with operating in this space.

Pressing issues include considerations of how to widen access to the Library's digital collections and the UK Web Archive in a legal framework that prioritises access conditions best suited to print. The Library is also grappling with digital preservation issues regarding how best to ingest, safeguard, and provide access to structured datasets such as maps, eBooks, interactive fiction, and new and evolving media.

This chapter outlines the growing digital expertise within the UK's Legal Deposit Libraries, and demonstrates the agile approach that is necessary to ensure that the current, ever changing UK publishing environment is safeguarded and available for current and future generations to consult.

Legal Deposit Libraries Act 2003, http://www.legislation.gov.uk/ukpga/2003/28/contents

² Legal Deposit Libraries (Non-Print Works) Regulations 2013, http://www.legislation.gov.uk/ukdsi/2013/9780111533703/contents

1. Context and history

The legal deposit collections of the National Library of Scotland have their foundations in The Advocates Library of the Faculty of Advocates.³⁴ The Advocates Library was formally inaugurated in 1689, although the Faculty had long been collecting texts and other documents to support legal practice.⁵ The Copyright Act of 1710⁶ entitled the Keeper of the Advocates Library to claim a copy of every book published in the British Isles, the first form of legal deposit, thereby broadening the scope of the Faculty's collection beyond legal materials. As a result of the Act, The Advocates Library became Scotland's national library in all but name. By the beginning of the 20th century the ever-growing collection of legal deposit materials was proving challenging for the Faculty, a private members' organisation, to store and maintain. In seeking a solution it approached the UK government for support. In 1925 the National Library of Scotland Act⁷ established a new library for Scotland, and the Faculty gifted 750,000 non-legal books, pamphlets, maps and sheet music as the founding collection for the National Library of Scotland. With the Act, responsibility for legal deposit transferred to the new National Library, although the Faculty of Advocates retained the right of legal deposit for legal works.

To store and manage the national collection a new building was erected in Edinburgh's city centre, adjacent to the Faculty, opening in 1956. With the ever-increasing output from UK publishers, the Library soon outgrew these premises, and in 1989 a further building, chiefly for the storage of materials, opened at Causewayside, a mile from the city centre.⁸

In 2016 a third facility opened at Kelvin Hall⁹ in Glasgow, providing access to the Library's non-print legal deposit, digitised,¹⁰ and moving image¹¹¹² and sound collections.¹³ The centre at Kelvin Hall was specifically designed to be a welcoming, comfortable and open space, where anyone may enter and freely view, use and enjoy the collections. Unlike the Library's centres in Edinburgh, where membership is compulsory before physical and non-print legal deposit collections may be consulted, Kelvin Hall has no such restrictions. This openness, along with the mix of interactive services for digital and moving image collections, has helped challenge the

³ The Faculty of Advocates is a Scottish body of independent lawyers, http://www.advocates.org.uk/

⁴ Solberg, Thorvald. "History of the Legal Deposit of Books throughout the British Empire." The Library Quarterly: Information, Community, Policy, vol. 8, no. 4, 1938, pp. 528–536. JSTOR, www.jstor.org/stable/4302514.

⁵ The Advocates Library http://www.advocates.org.uk/faculty-of-advocates/the-advocates-library

⁶ Copyright Act, 1710 http://www.copyrighthistory.com/anne5.html

⁷ National Library of Scotland Act 1925 (repealed) https://www.legislation.gov.uk/ukpga/Geo5/15-16/73/contents

⁸ Brief history of the National Library https://www.nls.uk/about-us/what-we-are/history

⁹ National Library of Scotland at Kelvin Hall https://www.nls.uk/kelvin-hall

¹⁰ Digitised collections, National Library of Scotland https://digital.nls.uk

¹¹ Moving image collections, National Library of Scotland https://www.nls.uk/collections/moving-image-archive

¹² Moving Image Archive at Kelvin Hall, National Library of Scotland https://www.nls.uk/collections/moving-image-archive/visit

¹³ Sound collections, National Library of Scotland https://scotlands-sounds.nls.uk/

perception that the Library is an austere, serious research institution, but rather the Library is for anybody and has something for everyone.

2. National Library of Scotland and the changing publishing environment

In the 1990s, with the proliferation of low-cost computing and the web as a new, open and democratising platform, publishing underwent a rapid and unforeseen expansion. Works were published in both print and digital formats simultaneously, born-digital works emerged, and a new publishing paradigm arose where potentially anyone could be a publisher. This shift, from print to digital, and the dissemination of digital works outwith traditional publishing routes challenged the Library's ability to comprehensively collect Scotland's published heritage¹⁴. Along with the other legal deposit libraries and interested parties, the Library lobbied government and publishers to extend legal deposit to include digital works ¹⁵ ¹⁶ . In 1997 the Secretary of State for Culture, Media and Sport (DCMS) set up a working party to advise on how a national archive of non-print content might be achieved. At the end of 1998, following consultation¹⁷, the Government accepted the recommendations of the working party, namely that only statutory deposit could secure a comprehensive national published archive¹⁸. The Secretary for State recommended that a technical group be set up to look at definitions and investigate impact on business, before drafting legislation, and in the meantime that a voluntary code be developed, agreed between publishers and deposit libraries. This led to the development of the Code of Practice for Voluntary Deposit of Non-print Publications, which was adopted at the end of 1999¹⁹.

Under the auspices of the voluntary code, and prior to the introduction of non-print legal deposit regulations in 2013, the Library and other legal deposit libraries were able to capture elements of the UK's growing digital publishing outputs. Over this period, the Library negotiated agreements ²⁰with many Scottish publishers to ensure that their born-digital publications would be deposited with the Library. In a parallel development in 2003, the UK Web Archiving Consortium (UKWAC)²¹ was established, with the Library a founding member ²². This followed

https://www.webarchive.org.uk/wayback/en/archive/20130424101504/http://www.bl.uk/aboutus/stratpolprog/legaldep/voluntarydeposit/ (UK Web Archive, open access)

¹⁴ Conference of Directors of National Libraries Working Group. The legal deposit of electronic publications, UNESCO, 1996. Available at https://unesdoc.unesco.org/ark:/48223/pf0000105504 (UNESCO website, open access)

¹⁵ British Library (1996) Proposal for the legal deposit of non-print publications to the Department of National Heritage from the British library. London: British Library, 1996.

¹⁶ Clive Field, Securing digital deposit in the UK: the Legal Deposit Libraries Act 2003', IN Alexandria, 16.2 (2004), 87-111, available at: https://doi.org/10.1177/095574900401600204 (Sage journals, login or purchase required)

¹⁷ Department for Culture, Media and Sport (1998) Legal deposit of publications: summary of responses to the consultation paper Issued by the Department of National Heritage, the Scottish Office, the Welsh Office and the Department of Education Northern Ireland. London: DCMS, 1998.

¹⁸ House of Commons (1998) 'Publications.' Hansard written answer,17 December 1998, pt. 43, col. 682 ¹⁹ Working Party on Legal Deposit (1999) Code of practice for the voluntary deposit of non-print publications. Online transcription available at:

²⁰ Recorded within the Library's Legal Deposit Claims & Receipts Database.

²¹ UK Web Archiving Consortium https://en.wikipedia.org/wiki/UK_Web_Archiving_Consortium

a similar voluntary, permission-based approach to collecting digital content, and enabled the Library to harvest Scottish websites and online documents, as well as to test some of the requirements of future legal deposit regulation. Where feasible, agreements also gave permission for these resources to be made available openly. Publishers that entered into these agreements included the Scottish Parliament and other public sector bodies, non-commercial bodies such as the Society of Antiquaries of Scotland, website owners and private individuals. These voluntary deposit and web harvesting approaches led, in 2009, to a memorandum with the Scottish Government for collecting its digital content ²⁴. The Scottish Government had by this time switched the majority of its publishing output to digital format, with few publications produced in print, making it essential that the digital copies were retained for posterity in a trusted repository.

With the introduction of the 2003 Legal Deposit Act²⁵ non-print publications were incorporated within the UK legal deposit framework, but as enabling legislation there was much work still to be done before non-print legal deposit would become a reality. Research was carried over this period to support the development of the regulations ²⁶²⁷²⁸²⁹ and an Advisory Panel was set up to manage implementation³⁰.

²² UK Web Archive Consortium members

https://web.archive.org/web/20100807233342/http://www.webarchive.org.uk/ukwa/info/about

- ²³ Society of Antiquaries of Scotland https://www.socantscot.org/
- ²⁴ Memorandum of understanding for the collection and preservation of Scottish Government web content ... between National Library of Scotland and Scottish Government, 1st July 2009. (Internal document)
- ²⁵ Legal Deposit Libraries Act 2003, http://www.legislation.gov.uk/ukpga/2003/28/contents
- ²⁶ Michael Day, Collecting and preserving the World Wide Web: a feasibility study for the JISC and the Wellcome Trust, Bath: UKOLN for Wellcome Trust and JISC, 2003. Archived copy available at: https://web.archive.org/web/20041028035721/http://library.wellcome.ac.uk/assets/WTL039229.pdf (Internet Archive, open access)
- ²⁷ Legal issues relating to the archiving of Internet resources in the UK, EU, USA and Australia: a study undertaken for the JISC and Wellcome Trust, by Andrew Charlesworth (Centre for IT and Law, University of Bristol), 2003. Archived copy available at:
- https://web.archive.org/web/20040704110603/http://library.wellcome.ac.uk/assets/WTL039230.pdf (Internet Archive, open access)
- ²⁸ David Powel. Refining the map of the universe of electronic publications potentially eligible for legal deposit: updated final report: prepared for the Legal Deposit Advisory Panel. Electronic Publishing Services, (2006). Archived copy available at:
- https://web.archive.org/web/20080727024130/http://www.culture.gov.uk/images/publications/EPS_Report_to_LDAP_Nov_2006.pdf (Internet Archive, open access)
- Paul Cunnea. 'Selective web archiving in the UK: a perspective of the National Library of Scotland within UK Web Archiving Consortium (UKWAC)' IN SCONUL Focus Number 34 (Spring) 2005, pp 44-49. Available at: https://www.sconul.ac.uk/sites/default/files/documents/SCONUL%20Focus%2034.pdf (accessed Oct 2019)
- ³⁰ Department of Culture, Media and Sport. Legal Deposit Advisory Panel terms of reference, 2007. Archived copy available at:
- https://web.archive.org/web/20080727030355/http://www.culture.gov.uk/images/publications/LDAPToRMarch2007.pdf (Internet Archive, open access)

From 2008 the British Library led a collaborative project ³¹ with the National Library of Scotland and National Library of Wales to build a digital library system to support ingest of non-print works and their metadata. With the passing of the 2013 Regulations ³², and the resultant legal duty and right to collect non-print works, this technical infrastructure was used by legal deposit libraries to manage and provide access to non-print works. The core facilities of this infrastructure are:

Storage and security: For print materials security is achieved by each legal deposit library receiving a physical copy and storing it in temperature-controlled, secure buildings. For non-print materials the British Library coordinates ingest of most digital resources into the shared digital library system. The system then distributes the resources to four digital storage "nodes" at the British Library at St Pancras, the British Library at Boston Spa, and the national libraries of Scotland and Wales.³³ If a resource on a node becomes damaged, for example through bitrot, or is lost as a result of a disk failure, the node attempts to resolve the issue from its local disk array. However, if it is unable to do this it will then seek a replacement copy from another node. For digital deposit, a single copy at a single point of deposit – as opposed to 6 physical copies for each of the six legal deposit libraries - leads to a more efficient deposit method for both publishers and libraries, and ensures that all six legal deposit libraries have access to the same corpus of legal deposit material; this would not necessarily by the case for print deposit, e.g. failure of a publisher to deposit or the libraries to request six print copies.

Access: The Library's access to non-print legal deposit resources is provided via Library Search,³⁴ a cloud-based resource discovery system that integrates all of the Library's collections whether physical or digital. For the vast majority of non-print legal deposit works, the technology behind Library Search harvests new and updated metadata from the British Library metadata aggregation service. Library Search then processes and indexes this metadata to make it discoverable.

When a user finds a non-print legal deposit work in Library Search there are several layers of control that moderate access and use of the resource, in accordance with the criteria set out in Part 4 of the 2013 Regulations:

the user must be on the Library's premises and using a Library provided "terminal"³⁵; before using a resource they must accept the terms and conditions of use; they may not annotate or take copies of the resource, although they may ask Library staff to print limited copies within what is allowed under the regulation.

³¹ Memorandum of Understanding between the British Library Board, The National Library of Scotland, and the National Library of Wales for Shared Technical Infrastructure for Managing UK Electronic Legal Deposit Materials, 2009 (internal document)

³² Legal Deposit Libraries (Non-Print Works) Regulations 2013, http://www.legislation.gov.uk/ukdsi/2013/9780111533703/contents

³³ Security for NPLD publications https://www.bl.uk/legal-deposit/security-for-publications

³⁴ Library Search, National Library of Scotland https://search.nls.uk/

³⁵ The Legal Deposit Libraries (Non-Print Works) Regulations 2013, Reader access to relevant material, paragraph 23 https://www.legislation.gov.uk/ukdsi/2013/9780111533703/part/4

- Library Search NPLD eBook
- NPLD terms and conditions
- NPLD ebook viewer
- NPLD article viewer

This is implemented in the Library by providing users with PCs in the reading rooms in Edinburgh and at the Kelvin Hall centre in Glasgow. At the time of writing, users may not use their personal laptops or phones to access non-print legal deposit resources, and this causes them considerable inconvenience as they have to move from their desk where they are consulting physical resources to another location in the Library to consult the non-print legal deposit materials.

In the event of downtime on the Library's node, whether that be planned or unplanned, access to non-print legal deposit resources is maintained by "failing-over" to another node, usually the node at the British Library at St Pancras in London. That node then provides access to Library users until the downtime has passed and access restored from the local node. This system has been in place since 2016 and is used on average three to four times a year, most often to support ongoing access during system maintenance.

This infrastructure has enabled the Legal Deposit Libraries to safeguard and provide access to a growing collection of digital material since the non-print regulations came into force, which complement the print holdings. At time of writing the National Library of Scotland's collection is in excess of 52 million resources with approximately 26 million physical items and 45 million digital items. The 27 million digital items are a combination of non-print legal deposit resources, collections digitised by the Library and digital collections purchased, acquired or donated to the Library. The legal deposit collections of the Library are made up of approximately 13 million print items collected since the Copyright Act, 1710 and more than 16 million non-print legal deposit items, of which approximately 6.9 million are ebooks and articles³⁶ and 10 million are archived websites, collected since the 2013 non-print regulation^{37 38}.

Receipt and ingest of digital resources is now a core activity for the Library, which brings with it various challenges, opportunities, and issues, given the legal frameworks within which the Library operates. The first set of these challenges can be framed as Challenges of Access: how can the National Library of Scotland allow users to access the growing digital collections in ways which respond to their affordances within the legal framework? The second set of challenges can be framed as Challenges of Format: how can the Library effectively deal with ever-changing and developing digital formats, in ways which will allow them to be used, adapted, searched, safeguarded, and preserved for future generations?

³⁶ Library Search search for non-print legal deposit articles and ebooks https://search.nls.uk/primo-explore/search?query=any,contains,44NLS_NPLD&tab=tab1&search_scope=TAB1_SCOPE1&vid=44NLS_NU1

³⁷ National Library of Scotland. Annual Review 2019/19. Available at: https://www.nls.uk/about-us/publications/annual-review/2018-2019

³⁸ Annual Acquisitions and Projected Intake (Internal statistics, unpublished)

3. Challenges of access

3.1 Widening access to the National of Library of Scotland's collections

Since 2004, successive National Library of Scotland strategies have made an increasingly strong commitment to widening access to the national collections, and collecting digital content and providing digital copies of its physical collections is a core manifestation of that commitment. The current strategy, *The Way Forward 2015-2020*,³⁹ is the most ambitious yet, with a key objective stating that, by the Library's centenary in 2025, it will have further widened access to its collections by having one third of its holdings in digital format. This ambitious figure (around 15 million on current projections) will include both physical items from the collection that have been digitised, and the non-print legal deposit collection. At the time of writing (2019), the Library is over half way towards this target.

Over the years this drive to widen access by providing digital copies has taken different approaches, but always with the same aim. These methods have included:

The digitisation of physical collections. For many years there has been an ongoing programme to digitise physical items, such as books, maps, photographs, posters, and music. Once digitised, these resources are made available for free globally and, where rights and permissions permit, as open access. For example, the Library holds one of the largest and most comprehensive collections of maps in the world⁴⁰ and since 2010 a digitisation program has made available online more than 200,000 maps to a wide global audience via a popular digital mapping service.⁴¹⁴² Other examples include:

- Medical History of British India ⁴³. This digitised collection, taken from the Library's India Papers Collection, provides access to rare documents and entry into the history of disease and its prevention in 19th and 20th-century British India, between 1850 and 1950.
- Scottish Post Office Directories⁴⁴ With over 700 digitised directories covering most of Scotland and dating from 1773 to 1911, this collection provides the an excellent basis for researching Scotland's family, trade, and town history, enabling users to easily find out where people lived at a certain time and how they earned their living
- The Word on the Street ⁴⁵ and the Broadsides from the Crawford Collection ⁴⁶. These two collections of amounting to around 4,000 broadsides from 1501 to 1910 provide a

³⁹ https://www.nls.uk/media/1194639/2015-strategy-2.pdf

⁴⁰ National Library of Scotland map collections https://www.nls.uk/collections/maps

⁴¹ National Library of Scotland digital maps service https://maps.nls.uk/

⁴² maps.nls.uk Google Analytics Apr 2018-Mar 2019; 1.32 million users, 3.33 million view sessions, 33.3 million page views

⁴³ Medical History of British India. Available at: https://digital.nls.uk/indiapapers/

⁴⁴ Scottish Post Office Directories Available at: https://digital.nls.uk/directories/

⁴⁵ The Word on the Street. Available at: https://digital.nls.uk/broadsides/

⁴⁶ Broadsides from the Crawford Collection. Available at: https://digital.nls.uk/broadsides-from-the-crawford-collection/archive/144782273

- fascinating insight into Scotland over this period, covering crime, politics, religion, romance, emigration, humour, tragedy, royalty and superstitions.
- British Military Lists ⁴⁷ and the Rolls of Honour ⁴⁸. These two collections provide details
 of officers who served in the British Army, Navy and Air Force in the First and Second
 World War, plus Scottish listings of casualties and those who died on active service in
 the First World War, provide an invaluable resource for researchers and family
 historians.

The digitisation programme was expanded in 2018, adding more digitisation stations and staff, with the goal of significantly increasing digitisation throughput.

Digital deposit, harvesting and donation. As noted in the previous section, following on from the earlier voluntary deposit and permissions based access agreements, the Library has continued to seek wider access agreements for digital content and websites collected under legal deposit, where this is feasible and resources allow. Where we have the necessary agreements in place, we make the content openly accessible online.

Purchased and subscription digital resources ⁴⁹. When purchasing or subscribing to digital content, the Library negotiates agreements with publishers that determine terms of use and access. To date we have negotiated remote access to over 80% of these collections, which enables Library users who are resident in Scotland to access the content from anywhere⁵⁰. This is a potentially transformative access to knowledge for readers, allowing citizens across Scotland remote, free access to resources as diverse as:

- 17th and 18th Century Burney Collection Newspapers full-text digital archive of 17 and 18th century newspapers, broadsides and news pamphlets
- Eighteenth Century Collections Online full-text digital archive of over 200,000 publications printed between 1701-1800.
- 19th Century UK Periodicals 600 full-text titles from the Library's collections.
- Access U.K. & Scotland Newspapers Full-text current content of local and regional newspapers, from over 860 titles.
- Archives of Sexuality & Gender: LGBTQ History and Culture Since 1940 consisting of 20 individual collections spanning 1940 to 2014 and amounting to around 1.5 million pages.
- British Library Newspapers, Parts 1 & 2 full-text digital archive of 19th century newspapers, with the complete runs of 70 national, regional and rural daily and weekly publications originating in England, Scotland, Ireland, and Wales
- Making of Modern World 12M pages on development of modern business and economics, from 1450-1850.

⁴⁷ British Military Lists. Available at: https://digital.nls.uk/british-military-lists/archive/97343435

⁴⁸ Rolls of Honour. Available at: https://digital.nls.uk/rolls-of-honour/archive/100261716

⁴⁹ National Library of Scotland. eResources. Available at: https://auth.nls.uk/eresources/

⁵⁰ Gooding, P. (2017) Historic Newspapers in the Digital Age: 'Search All About It', Routledge.

- Cold War Intelligence full-text collection of 2,360 formerly classified U.S. government documents detailing the successes and failures of the U.S. intelligence community in its efforts to spy on the Soviet Union during the Cold War.
- Scotsman Digital Archive every issue of the iconic newspaper from 1817 to 1950.
- PerditaManuscripts full-text collection of early modern women authors from Britain, 1500-1700.
- British & Irish Women's Letter & Diaries 100,000 pages of published letters and diaries, spanning 1500 to 1950.
- D&B Hoovers a business information database which provides in depth profiles of companies worldwide, industry profiles, news and market research
- Scientific American Archive full-text digital archive, covering all 4,960 issues from 1845 to 2005, and containing more than 133,700 articles.
- SpringerLink 2,000 journal titles and over 65,000 e-books
- Factiva contents of over 14,000 leading newspaper titles, news sources and company information from around the world
- Electronic Enlightenment over 70,000 letters and documents from over 8,500 correspondents, from 1609-1900.
- Web of Science citation database covering over 10,000 high-impact journals in the sciences, social sciences, arts and humanities.

Establishing Kelvin Hall in Glasgow as a 'digital access point'. With the opening of the Kelvin Hall facility in 2016, we not only provided a physical presence for the National Library in the centre of Glasgow, and a suitable, modern facility for consulting the moving image and sound collections, we also provided an additional location for accessing the Library's digital collections, specifically where access is restricted to Library premises. This means we can now provide access from a new geographic location to the 20% of purchased and subscribed digital content that can only be viewed on Library premises due to the licencing conditions noted above. This allows us to deliver these valuable resources to the largest population centre within Scotland, where terms and conditions allow. This also has major implications for increasing access to NPLD restricted content. Following discussion at JCLD, and with agreement from the publisher representatives, it was accepted that the regulation provided for access to NPLD content from the Library's new Kelvin Hall premises. This will extend the availability and accessibility of the collections significantly to the people of Scotland, as well as the services provided at the Kelvin hall facility

With the advent of non-print legal deposit regulations in 20013, after many year's preparation, the potential for increasing the Libraries digital holdings further was enormous. However, the extent to which this would help increase access to the Library's collections for our users was not straightforward, despite some of the developments described above.

3.2 Restrictions on Access, Use and Reuse of NPLD content

The introduction of the 2013 Regulations has greatly advanced the Library's strategic objective of holding a third of its collections digitally. In the first few years of NPLD, many large publishers have "transitioned" from print to non-print legal deposit and the legal deposit libraries' shared

web harvesting programme has expanded to include annual snapshots of what is now more than 10 million websites from the open UK web domain. However, access and use restrictions in the 2013 Regulations do not align well with the Library's wider access ambitions, and threaten to restrict access to digital content, often despite the interests of many digital publishers, such as bloggers, private individuals and non-commercial organisations and publishers.

The principle objective of the 2013 Regulations, as stated in the 2012 final impact assessment⁵¹, is "to allow the preservation of non-print publications for future generations". In the recently published Post Implementation Review of the Regulations⁵², the Joint Committee on Legal Deposit (JCLD) considers three areas in its assessment of the success of this objective so far: collection; preservation; and access and use.

The review recognises that, five years after its introduction, implementation is an ongoing process, but suggests that collection and preservation can be considered to have been largely successful. Indeed, by year 4 (2016/17) of implementation collection was 23% over the highest estimates for year 5, with 58% of all monographs and 35% of all journals identified by the libraries now collected in digital rather than print format.

However, the review provides less agreement on the level of success in terms of access and use, with significantly different interpretations between the publisher and library representatives within JCLD. From its inception, a key requirement of the Regulations was to achieve a balance between the interests of publishers and rights holders, and the needs of libraries, researchers, and the wider public. As such, conditions controlling terms of access and use are central to the Regulations.

In line with this, the use of NPLD content is tightly controlled, with saving, downloading, and taking images of NPLD content prohibited. This is in part technically controlled using a locked down browser when accessing NPLD content. Printing is permissible within standard copyright restrictions, and is offered via the Library's networked printers. Data and text mining is also not permitted, restricting the ability of researchers to analyse UK publishing output and the development of the UK web, something which the growing digital corpus is ideally suited to meet, and which would be of value to a wide range of stakeholders, including the publishing industry.

In terms of accessibility, provision is made in the regulations for Libraries to provide accessible copies of relevant material for a visually impaired person. This is achieved by providing access to NPLD content via dedicated terminals with accessibility software installed. However, people with accessibility requirements are unable to use the accessibility tools and software that they

⁵¹ Department for Digital, Culture, Media & Sport (2012) Impact assessment for the Legal Deposit Libraries (Non-Print Works) Regulations 2013. Available at: https://www.legislation.gov.uk/ukdsi/2013/9780111533703/impacts

⁵² Department for Digital, Culture, Media & Sport (2019) Post-Implementation Review of the Legal Deposit Libraries (Non-Print Works) Regulations 2013.. Available at: https://www.gov.uk/government/publications/post-implementation-review-of-the-legal-deposit-libraries-non-print-works-regulations-2013

are familiar with, which can present a barrier to use. In addition, this provision is limited to people with visual impairments, whereas the current accessibility legislation⁵³ covers anyone with a disability. This deficit has been recognised by all parties in the Post Implementation Review, which recommends that the regulations are brought in line with current accessibility legislation.

The Regulations more fundamentally restrict access to NPLD materials to the premises of a legal deposit library. They also restrict access to only one copy of any one resource within a library at any one time – mirroring the print model. However, these restrictions are notably more restrictive than those that apply to print legal deposit and to other collection formats. As such, although the print environment may have formed the basis for formulating non-print legal deposit access conditions, the Regulations can be viewed as enforcing more stringent conditions than apply to print, as highlighted in the Digital Futures White Paper by Gooding et al

Compared to print, the digital domain is significantly more diverse and complex. There are a huge and varied number of digital content producers, content owners and publishers, the majority of whom have never produced print works. This diverse group is not represented by traditional publisher bodies, and will have a range of views on digital collecting, preservation and access potentially very different from traditional publishers, commercial and noncommercial.

For example, many scholarly publications published today are made available under open licences such as Creative Commons ⁵⁵ to promote open access and re-use. Furthermore, a large amount of public information is made available under Open Government Licencing⁵⁶, to encourage re-use, in accordance with the principles of the Re-use of Public Sector Information Regulations 2015⁵⁷ open data strategies ⁵⁸. And blog or social media content creators are more likely to be more interested in openly sharing their content, rather than restricting access.

However, the Regulations place the same access conditions across all content collected under non-print legal deposit, irrespective of the interests, wishes or culture of the content and rights owners. By placing blanket access and use restrictions on this material, legal deposit libraries are in effect limiting access to huge volumes of publicly accessible content. This is not an issue for researchers and the public while this content remains available on the open web, although it does limit the value and utility of the corpus of material, but it does present a problem when this content is no longer available.

⁵³ Equality Act 2010. Available at: http://www.legislation.gov.uk/ukpga/2010/15/contents

⁵⁴ Gooding, P., Terras, M. & Berube, L. (2019). Towards User-Centric Evaluation of UK Non-Print Legal Deposit: A Digital Library Futures White Paper. Available at: http://eprints.gla.ac.uk/186755/1/186755.pdf

⁵⁵ Creative Commons website. Available at: https://creativecommons.org/

⁵⁶ Open Government Licence for Public Bodies. Available at: http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/

http://www.nationalarchives.gov.uk/doc/open-government-licence/version for Re-use of Public Sector Information Regulations 2015. Available at:

http://www.legislation.gov.uk/uksi/2015/1415/contents/made

⁵⁸ Scottish Government Digital Directorate (2015) Open Data Strategy. Available at: https://www.gov.scot/publications/open-data-strategy/

This disparity in access between the originally published content, and the NPLD restricted content, is most clearly demonstrated with the UK Web Archive.

3.3 Access and the Web Archive

Analysis by Andy Jackson, UK Web Archive Technical Lead at the British Library, found that by 2015, 24 months after UK legal deposit web harvesting began, nearly 40% of the content harvested in 2013 had disappeared from the live web, was missing or had moved ⁵⁹. Indeed, deeper analysis found that link rot and content drift were the norm, and that 50% of web resources were gone or unrecognisable after 1 year, 60% after 2 years, and 65% after 3 years.

Looking further back, less than 20% of content archived from 2004 can be confirmed as identical to content currently on the open web. Bear in mind that, prior to the 2013 legal deposit regulations we were archiving a tiny fraction of the UK domain under permission-based harvesting, so the vast majority of UK web content will be lost unless captured by non-statutory web harvesting bodies, such as the Internet Archive ⁶⁰, who have been harvesting the web globally since 1996.

The analysis makes clear that a large proportion of even recent web history is either no longer extant, is difficult to find, or has changed out of recognition. Since the NPLD Regulations came into force in 2013 we are able to collect and preserve the UK web domain at scale. Much of the UK web domain from this period will be available for future generations to view and analyse due to this work, but for the time being much of this originally public content will remain behind closed doers, accessible only on legal deposit library premises.

This demonstrates how the Regulation, whilst empowering the libraries to make great strides in collecting and preserving digital content, at the same time limits access to once open, public content that is not readily available by other means. This presents a conundrum. The legal deposit libraries recognise that the interests of the traditional publishing industry require protection, and also that not all free-to-access content is necessarily non-commercial, for example models for monetising content through click-throughs and advertising, However the Regulations' single approach to access, which gives no provision for granularity in access conditions or publishing models and cultures, is at variance with the intentions of many digital content producers.

This conundrum can be further demonstrated by the options open to the libraries to make archived web content more accessible. The archive consists of an annual snapshot or "harvest" of the open UK web domain; this excludes any content that lies behind a paywall, or requires membership or a login to access. This annual snapshot is further supplemented by "selective harvesting". These selective harvests may focus on a particular event or topic, such as the 2005

⁵⁹ Andy Jackson (2015) Ten years of the UK Web Archive: what have we saved? Available at: http://blogs.bl.uk/webarchive/2015/09/ten-years-of-the-uk-web-archive-what-have-we-saved.html
⁶⁰ Internet Archive. Available at: https://archive.org/

G8 Summit in Scotland ⁶¹or the Scottish Independence Referendum in 2014⁶², and harvesting of this content may be carried out more frequently, have higher quality assurance measures, better description, and - of relevance to this section - more resources assigned to pursue open access agreements.

As noted above, since the introduction of the 2013 Regulations the National Library of Scotland has continued to pursue wider access agreements with website owners and non-commercial publishers. Until recently, this has involved collecting a second copy, in addition to the legal deposit copy, and making it globally available, with the rights holder's permission, either on the open UK web archive or via the National Library's web services. Over time, and working with the other deposit libraries, this approach has been standardised, and integrated into the non-print legal deposit technical infrastructure, specifically for web and document harvesting. In this context, and where wider-access has been agreed, access restrictions may be relaxed to allow the legal deposit copy to be openly and globally available.

This approach now applies to content where there is a direct agreement with a website owner or publisher, and through recognised open access agreements, such as the Open Government Licence⁶³ and the Library's memorandum with the Scottish Government⁶⁴.

However, adoption of open licences by public bodies is patchy, and despite open access research mandates and campaigns it is still far from being universal in the academic sector. In the meantime, pursuing individual agreements with content producers is resource intensive, and coupled with a low response rate means that the percentage of the web archive that is openly available using this approach will remain extremely low under the current framework. Taking the UK Web Archive as a whole, less than 1% of the archived content is globally accessible, despite all content being harvested from the from the open UK web domain.

This impact of restricting access is further highlighted with the merging of the Open UK Web Archive (containing permission-based content harvested since 2004) and the UK Legal Deposit Web Archive through a single search interface⁶⁵. This recent development integrates discovery of the openly accessible archived website and those that can only be viewed by visiting deposit library premises. The single search interface is a positive development, with restricted access sites now more visible, and sitting alongside open access archived sites. This raises the visibility of legal deposit corpus, helping to promote the archive to researchers and library users, making it easier to view search results and identity what has and hasn't been archived, even if

⁶¹ G8 Summit 2005 UK web archive collection https://www.webarchive.org.uk/en/ukwa/collection/2398

⁶² Scottish Independence Referendum UK web archive collection, https://www.webarchive.org.uk/en/ukwa/collection/65

⁶³ Open Government Licence for Public Bodies. Available at: http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/

⁶⁴ Memorandum of understanding for the collection and preservation of Scottish Government web content ... between National Library of Scotland and Scottish Government, 1st July 2009. (Internal document)

⁶⁵ https://www.webarchive.org.uk/

in most cases it will require travel to legal deposit library to view the archived websites themselves.

However, levels of use of the content based on access level, whilst not overly surprising, highlight and reinforce the impact restricted access has on this increasingly valuable dataset. As might be expected, the level of use and size of the two archive components - open and restricted - are inversely proportional, with the much smaller open access component (0.15% in 2017) having over 100 x the level of use than the exponentially bigger restricted access component (table 1). This severely hampers the promotion of the archive to researchers and the public in general and selling its inherent value to the research community.

	Sites	ze Data	Uses 2016/17	Page views 2016/17
Legal Deposit Uk Web Archive (access on LDL premises)	9-10 million	470 TB	2,311	16,036
Permissions cleared UK Web Archive (access on open web)	15,000	30TB	292,699	1,070,160

Table 1. Comparison of usage and views of Legal deposit versus permission-cleared web archives 66

In addition to these limitations on access, the restrictions on use note previously also severely limit the intrinsic value of the dataset for researchers. What could be an incredibly rich resource using modern text and data mining techniques for investigating, for example, social, linguistic, technical, and publishing trends, is instead limited to a search and discovery platform for the UK web over time. Valuable though this is, the research opportunity for mining at scale such an extensive dataset, with its comprehensiveness, span, diversity, cohesion and set boundary, is lost - at least, for the time being.

3.4 Future of Access & Use of NPLD

The JCLD review highlighted the tension between protecting publisher interests and providing public access to NPLD, as well as the diverging views of publisher and library representatives on JCLD in these areas. On the one hand, the libraries assert the value of legal deposit in terms of access and use, advocating in particular for a change to regulations to address the issue of access to the web archive in particular. Publisher representatives, on the other hand, question the role of legal deposit beyond collection and preservation for future generations.

Despite the differences in perspective between JCLD members, there appears to be some agreement on how to resolve these differences. While questioning the extent to which access and use is a key objective of legal deposit regulations, JCLD publisher representatives agree that a public consultation on access of the UK web archive may be a way forward. However,

⁶⁶ Post Implementation Review (Annex_A, Libraries_Submission) https://www.gov.uk/government/publications/post-implementation-review-of-the-legal-deposit-libraries-non-print-works-regulations-2013

publisher and library representatives agree that if this is pursued, a framework for protecting the interests of publishers and other rights holders would be required.

For open access advocates this may appear an over-cautious approach. However, there are important considerations to take into account. These include not just commercial interests, but also the moral and legal interests of rights holders in general. Any changes to the regulations need to consider what safeguards would be required, for example to protect future, unforeseen revenue streams. Legal deposit libraries will continue to have a responsibility to ensure the interests of rights holders are protected, and that legal deposit collections are used responsibly, while also meeting the needs of library users, researchers and the wider public.

UK non-print legal deposit is still in the early stages of implementation. There is an opportunity that a more balanced access framework can be established which not only recognises the interests of publishers and rights holders, and the diverse range of content producers and publishers across the UK digital published and web domain, but also the value that wider access to our digital and web heritage can bring. Such a framework needs to acknowledge that legal deposit libraries are collecting for current as well as future generations, and that within the digital publishing domain there are many more models and interests to consider and adapt to than in the print environment; one size is unlikely to fit all. If an effective balance is struck between the varying interests and needs of publishers, content producers, rights holders, researchers and the public, a non-print legal deposit access framework may emerge for the UK that is the envy of the world.

4. Challenges of format

As well as issues of access, non-print legal deposit highlights issues of changing digital formats, and how libraries can deal with a fluctuating and changing digital environment as well as the inherent tension in trying to make often ephemeral formats available for posterity. The 2013 Regulations are based on the print world. As such, the UK's legal deposit legislation focuses on the collection of a comparatively traditional understanding of electronic works. Legal deposit libraries are entitled to collect 'copies' of works that are published in non-print formats, whether on- or off-line. Focus on works distributed by 'copies' is historically and legally logicial, but not necessarily best-suited to the realities of digital content. To understand why this approach may not be the most appropriate for digital works, it is useful to reflect on the historic logic and legal context of collecting 'copies' under legal deposit.

The historic logic for 'copies' reflects the reality found in monographic and continuous print publications: when a monographic work is prepared for print publication, its content becomes complete, fixed and static at point of printing. This is true regardless of whether the print run is one copy or many copies; continuous works, such as newspapers, journals and magazines, are somewhat different. Individual issues are of course fixed and static copies at the point of printing, but the work itself continues to change through time as more issues are added. As a result, each time a new issue is printed, an extension or continuation of the work is produced.

The legal logic for 'copies' has its roots in copyright legislation. Under the UK's copyright laws, the right to 'issue copies' of a work to the public is the first of six rights reserved to the copyright owner.⁶⁷ Thus, the very notion of what it means to create a work worthy of copyright protection has been rooted in the distribution copies. Non-print legal deposit regulations continues this historical and legal context of copies and this logic gives the framework tangibility. Works are created, copies made available to the public in accordance with a copyright-owner's right, and a copy is duly collectable by legal deposit libraries.

The logic relating to 'copies' is suitable and appropriate for some digital content because, in many instances, digital works like ebooks and electronic journal articles and issues mimic print works, in that at point of production they become fixed, complete and static. However, there are many forms of digital content collectable under the 2013 Regulations for which this simple 'copies' logic is misaligned. In relation to the legislation, we refer to these as 'challenging formats'.

4.1 Challenging formats

The nature of digital works can be, and often is, radically distinct from the world of print. Consider a database or a website that is continually updated. When is the legal deposit 'copy' to be created or collected? Or a politician's social media feed where comments are added and deleted. When is the legal deposit 'copy' to be created or collected here? And what about a publication that is designed to 'decay' each time it is opened. How is this work preserved by legal deposit?

The nature of the digital domain, and its apparent infinite flexibility and adaptability to create new digital presentations, new formats and content, will continue to pose fresh legal deposit challenges. Who can imagine what a future digital work will look like or function? To focus our discussion, three specific types of challenging formats are explored: structured datasets, ebooks as apps, and interactive fiction.

4.1.1 Structured datasets

Substantial datasets, such as business and cartographic databases like those from the UK's mapping agency Ordnance Survey, can be considered challenging formats in the non-print legal deposit context because of their scale and their continuously-developing nature. Of course, there are certain types of print works that are similar to these datasets, and they have their own collecting challenges; consider 'updating services', which require loose leaf paper updates to be inserted into a work, replacing outdated leaves of information, which is removed. However, contemporary non-print equivalents present perhaps even more vexing challenges.

The very value of these works is vested in their size and currency. These are works that are made to be remade, to be continually updated and revised with the most accurate and current

-

⁶⁷ CDPA s.16 as amended

information. In the age of print this meant frequent distribution of loose leaf updates. With non-print works, this means databases, normally obtained through subscription, that are updated at intervals. Some data points may be updated numerous times a day. Others may not be updated for years, or at all. But the publication remains as one single 'live' database of the most current data - immeasurably useful to the immediate user, but immensely challenging to the logic of legal deposit and preservation of 'copies' of knowledge.

When is the legal deposit 'copy' to be obtained? Should a fresh copy of a vast dataset be collected afresh every time a datapoint is revised? Does the adjustment of a single data point mean that the work in its entirety is a 'new' version? Should the work be collected at fixed times, for example annually?

The challenge in terms of use is obvious when considering the future user. In ten years time, how will a legal deposit library provide access to the version of this database as it was on 1 January 2019? Not without some logistical effort, perhaps, that same library may be able to provide access to the database as it was on 1 January 1990, by pulling together the correct sequence of loose sheets. There is no ready solution, bar the creation of an unwieldy multiplicity of copies of each database, to ensuring that the database of 1 January 2019 is as accessible to the future user. Consider computer-assisted legal research services that maintain databases of statutes, court documents and judgements that document the evolution of law and regulation. If a library is unable to take continuous copies, or overwrites an existing copy of such a database with the most current version, then what does this mean for a legal professional or court that is trying to understand the legal guidance or practice at a historical point in time? The evidence trail may be lost or incomplete, perhaps with profound consequences.

4.1.2 eBooks as apps

While books and ebooks are not generally challenging works, books published digitally as apps do pose unique challenges. By 'ebooks as apps' we mean digitally-published books that are published along with supporting content, or as part of a richer, content-heavy application. For example, ebook apps that not only contain text but also audio or rich imagery.

Unlike the scale and speed-of-change challenges of structured databases, it is the richness and presentation style of 'ebooks as apps' that are challenging in the non-print legal deposit context. These works are most often created in exceptional detail, with the traditional book text woven through and around other content, from music to video to imagery, that supplements and expands on the direct text.

While these works may be more straightforward to collect as 'copies', ensuring suitable access to and engagement with these works presents real hurdles. eBooks as apps are likely to be developed specifically for one particular consumption format. How these works can be properly accessed by current and future users of non-print legal deposit collections remains a challenge, as rarely are such works designed for indoor, seated, desktop terminal access.

4.1.3 Interactive fiction

In many respects, interactive fiction manifests the genius of digital. Works, whether or not based on text, that contain geocoding to localise a narrative to a user's time and place, or narratives through which the user can navigate and develop their own plots or conclusions are examples of interactive fiction. These works make the very best use of 'digital' - unlike 'flat' ejournals and ebooks, interactive fiction grabs and runs with the beauty and power of digital, combining methods, formats, and data sources and enabling a shattering of a fourth wall between the creator and the consumer.

Under non-print legal deposit, these works are also some of the most challenging to collect, preserve and provide access to, as well as being at more risk of loss. The challenges of access are most acute and obvious: how, within the limited confines of non-print legal deposit access, is a user wholly and properly to seek to interact with and enjoy these works? By their nature, the user experience may be designed to be wholly different on each occasion, or in any given context, location, or scenario. The challenges of collection are similarly not insignificant. What is to be collected? A device that is pivotal to the experience? Whose version of the narrative?

4.1.4 Digital as artefact

New interactive qualities of digital literature, for example where the text of a work alters as the reader physical engages with it, challenges the perception that digital and print works are basically analogous. This is brought into sharp focus when considering how such works may be exhibited as examples of "electrified literature".

For the purposes of display and exhibition print works are placed in secure atmospherically-controlled cases protecting them from theft and damage, while simultaneously making them visible to visitors to the Library. Would it be strange to put a Kindle or iPad behind glass so as to show particular non-print legal deposit publications? When books are placed in cases, they "change" from literature to exhibit, and in doing so say something additional to the content of the item itself. Should ebooks and ejournals be treated in the same way? And if so, how are electronically powered devices to be physically displayed? Consider an exhibition taking place in 2030, should digital works published in 2019 be displayed on the hardware of the time? What does it actually mean to preserve an ebook over time as a piece of cultural heritage, or has the carrier become meaningless? And if so, what does this imply?

5. Non-print resources as an agent of change

With the introduction of the Regulations in 2013, the expansion of digitisation and the growing investment in purchased digital collections, the Library has, in the past ten years, seen an unparalleled growth in its digital collections. This growth comparatively dwarves the growth of the physical collections. The need to manage the expansion in digital collections has required a change to the Library's skills portfolio, and has led to investment in digital skills development.

retraining of staff and the creation of new positions. This shift has also required the Library to consider new approaches and processes.

Specifically, some "traditional" skills, techniques and processes have become redundant with the management of non-print legal deposit collections, while others are evolving, and some are new. For example

- The acquisitions role and function remains, with staff claiming digital publications that have fallen outwith the legal deposit libraries' mainstream non-print legal deposit collecting. With print collections staff would normally unbox and process (for example, by adding the stamp of ownership and moving items to temporary shelving) up to 3,000 new publications a week. These processes are redundant with non-print collections, freeing staff to carry out intellectual rather than physical work.
- The metadata description function remains essential with non-print legal deposit. However, given the volume of this material the Library is unable to provide the quality of description previously feasible with print collections through the Legal Deposit Libraries Shared Cataloguing Programme. The Library does not have the capacity to routinely create new descriptions or upgrade and enhance existing descriptions for non-print legal deposit resources that have been provided by publishers or, in the case of the UK Web Archive, information scraped from the homepages of websites. However, the scale of non-print collections brings considerable opportunity for developing new tools and techniques of description, through the use of "big data" analysis and machine learning techniques. There is also untapped potential in using "the crowd" as an intellectual source for description of non-print collections, whether via analysis of search activity or involvement in tagging, annotating etc. Collaborating in the design of such tools and services will require metadata staff at the Library and the other legal deposit libraries to develop their knowledge and skills in data processing and in developing networks of expertise and participation.
- The traditional role of the Library's Collection Support staff who place recently-received physical items on shelves is redundant in relation to non-print materials. The digital equivalent of this role is carried out by the Digital Department's technical infrastructure team which oversees the maintenance and support of the "node", for example by adding storage as required. Outwith planned upgrades and maintenance support, the "node" requires few hours of trained technician time per month, compared with many hours of ten Collection Support staff per month in respect of physical materials. However, as the print legal deposit intake declines and Collection Support colleagues spend less time placing books and journals on these shelves, they are able to spend more time on value-added duties, such as managing and optimising the Library's 130 miles of shelving, auditing and checking the collections, and processing requests for the Library's mass digitisation programme.

⁶⁸ Legal Deposit Libraries Shared Cataloguing Programme, The British Library https://www.bl.uk/bibliographic/clscp.html

- The systems librarian role has evolved from overseeing the library management and resource discovery systems to a more data-centric role, with added responsibility for processing and automating non-print legal deposit and other digital collection metadata.
- In 2013, in recognition of the challenges associated with ensuring continued access to its digital collections, the Library created the new specialised role of Digital Preservation Manager, and in 2019 added a further supporting position. These roles co-ordinate and champion digital preservation in the Library and are responsible for safeguarding the Library's digital culture. They provide advice on digital preservation issues and practices including to the UK and Ireland Legal Deposit Libraries to ensure digital legal deposit content is adequately cared for. Working in this field requires the development of new digital skills, expertise and workflows that support the Library and are regularly shared with others who have a growing realisation of the need for digital preservation.

The non-print legal deposit collections, along with other digital collections, have challenged the Library's 300 year-old print-based processes and practices. In just ten years it has become necessary rapidly to develop the digital skills of those staff working with and managing these resources. Digital collections are compelling the Library, and other libraries, to modernise, embrace the digital, and to look to new techniques and tools for processing these materials.

6. Conclusion

The nature of 'non-print' collections and non-print legal deposit legislation presents opportunities and challenges in considering access to, engagement with, and display of such works. Overarchingly, this is about the very nature of non-print collections, to expose and celebrate what makes digital works distinct from and similar to, print legal deposit works. However in many respects the Regulations so restrict and limit the access and use of non-print legal deposit collections that they are in fact inferior in these aspects to print collections. These boundaries and constraints need to be explored and pushed. The National Library of Scotland and other UK legal deposit libraries are open to further discussions and public consultation on this, in particular focusing on the access issues relating to the UK Web Archive and its alignment with Regulations and UK copyright law, and the challenges in the increasing complexity and sophistication in forms of content.

The development of non-print collecting has been of wider benefit and welcome to the Library and the legal deposit libraries, enabling collection growth in increasingly significant areas of digital publication and communication, and in the development of new and varied digital skills and competencies among staff. It is perhaps unsurprising that the legislative framework remains challenging; unlike its print counterpart, non-print legal deposit remains a new development, which is certain to continue to evolve and refine.