

Joint Collecting Framework for UK Legal Deposit, 2015-2020

1. Introduction

1.1. Legal Deposit. The Legal Deposit Libraries aim to ensure that the UK's published output is collected systematically and as comprehensively as possible. Traditionally this has involved printed works but, following the Legal Deposit Libraries (Non-Print Works) Regulations 2013, this is now complemented by UK digital publications ("non-print works"). The Libraries set out their first collecting plans for Non-Print Legal Deposit in 2013. This Framework updates this in the light of the experience we have gained since then.

1.2 Significance of digital. The continuing deposit of printed works, where no digital equivalent has been deposited, remains central to the 'national memory' and will still be managed by the Legal Deposit Libraries. However, because of the scale of the digital challenge and the fact that many print publishers also publish in digital formats, the present framework concentrates on electronic publications. This includes the transition, for many publishers, from a print-based operation to one where digital formats are also produced.

2. General Approach

Four key elements inform the Framework: the need to address digital works at risk of outright loss; the desirability of taking in digital works which offer to current readers clear benefits over print publications; acknowledgement that an exploratory approach is best suited to the high rate of change in the digital sector; and acknowledgement that there are serious financial constraints on the Libraries, not only because of public sector downsizing but because of the nature of the digital publications themselves.

2.1. Digital works at risk of outright loss. Because of the scale of digital publication the Libraries are phasing in their collecting of these materials. The Libraries seek to focus especially, but not exclusively, on digital works which are thought to be at risk of outright loss or change, for example the e-only books of small publishers, or websites in the UK domain. The Libraries also expect to develop better capability for coverage of news, official publications, sheet music, and digital mapping.

2.2. Digital works providing other forms of value. Because there are also advantages for readers and publishers if some kinds of digital publication, not necessarily at risk, are also collected at this time, the Framework takes a mixed, pragmatic approach, including these as appropriate. For example, where there is publishing in both formats, reading room delivery times of digital publications compared to their print equivalent are clearly far quicker. For publishers, delivering digital works is much cheaper than posting the printed equivalents. For the Libraries, it is expected that the transition from print to electronic deposit will reduce burdens on physical storage space for contemporary titles, many of which will be in regular demand by readers.

2.3 Exploratory approach necessary. One characteristic of digital publication is the high rate of innovation and change within the various forms. There is also the presence of emerging media which, though they are publications, do not fit into the traditional 'book, journal or newspaper' classification (for example, frequently updated online subscription services). While the Libraries do

not have the resources to commit to the wholesale collection of every variant and every new kind of media, we will explore what challenges the new media present and what options there are for safeguarding them for future generations. This would involve an assessment of the technical and human resources required to underwrite the process.

2.4. Constraints. In a continuing climate of financial constraints, including staffing, it is important to recognise that Legal Deposit activities carried out in the digital field do not automatically translate to savings for the Libraries. In fact, considerable investment in digital infrastructure and in staff has already been made and further investment will be needed for some time. Savings in physical storage space would not be made if the Libraries were to focus solely on the capture of e-only publications. For these reasons, the exploratory approach outlined above will be extended to questions of efficiency in the processing of material, for example reviewing the level of detail required in resource discovery for digital works.

3. The Collecting Framework, 2015-2020

3.1. The Libraries will focus especially, but not exclusively, on digital publications thought to be at most risk of outright loss or change, whilst also adding those which provide other forms of value. The Libraries will continue to work collaboratively with a wide cross-section of publishers, for example through the Joint Committee on Legal Deposit and individually. We will continue to work with publishers who publish digital and print versions in parallel, with print works continuing to be received until transition to e-deposit has been agreed.

3.2 The Libraries will be especially active in the following areas:

1. *Websites.* We will continue to collect comprehensively the UK's freely available websites and associated web content, and enhance our end-to-end web archiving capability:
 - a. Significantly increasing the number of UK sites beyond .uk , e.g. sites eligible for deposit within .com
 - b. Further developing expertise in understanding, explaining, and enabling use of the content of the UK Web Archive, through
 - i. Supporting research based on archived web resources
 - ii. Data analytics and visualisation
 - iii. Redeveloping the Open UK Web Archive, a window on the Legal Deposit UK Web Archive
 - iv. Enhancing reading room access to the Legal Deposit UK Web Archive, including making collections more visible, and better rendering and display of archived websites
 - v. Further curator and reading room training
 - vi. Wider engagement with stakeholders, including researchers
 - c. Develop the Document Harvester for collecting publications embedded in the open web. This will increase productivity in collecting and cataloguing individual digital documents distributed over the web and enhance both web archiving and digital deposit collecting streams
 - d. Explore the capture of eligible content which is technically unavailable to web harvesting technology, e.g. published content behind password protection, dynamically generated and database driven content
 - e. Develop focussed collections within the UK Web Archive, building on the existing event and time-based collections and not limited to topicality
2. *E-books.* We will further develop the collecting of e-books,
 - a. Prioritising
 - i. Identification and ingest of e-books only available in digital form, and which can be managed within current legal deposit capabilities
 - ii. Transition to e-deposit of publishers whose record of print deposit has been poor
 - iii. Transition to e-deposit of smaller publishers, including self-publishing, and within smaller presses:

1. Personal witness: autobiography, biography, family history, first-hand accounts
 2. UK local studies (history and geography, including digital mapping)
 3. Creative texts, including fiction, poetry, graphic novel
 - iv. Transition to e-deposit of publishers whose e-journals we have previously committed to collect electronically, in respect of publishers who produce both e-books and e-journals
- b. Liaising with aggregating platforms which distribute e-books in the small press categories above, to assist in their deposit at scale
 - c. Use of the Libraries' Document Harvester to increase ingest and management of e-books embedded on the open web
 - d. Increasing the number of e-books deposited via the Libraries' Publisher Deposit Facility ("the Publishers' Portal")
 - e. Exploring the challenges of more complex e-books, for example those which contain dynamic elements or difficult formats (including sheet music), and seeking, if possible, scalable solutions
 - f. Selectively acquiring (through purchase or donation) print versions of collected e-books, where a significant current or future reader demand has been demonstrated for the print version, e.g. where the print work is visually rich or precise in a way that cannot be rendered digitally, such as art books. In some cases, where there is an overwhelming argument to do so, we will continue to accept a publisher's output in print rather than digital.
3. *E-journals*. We will further develop the collecting of e-journals
 - a. Prioritising
 - i. Identification and ingest of e-journals only available in digital form, and which can be managed within current legal deposit capabilities
 - ii. Transitioning of publishers we have already entered into discussions with or committed to
 - b. Using the Libraries' Document Harvester to increase ingest and management of e-journals embedded on the open web
 - c. Exploring the challenges of more complex e-journals, for example those which contain separable chapters and dynamic elements, and seeking scalable solutions
 4. *Complex digital media*. We will explore the challenges of more complex digital media and take pragmatic decisions to maintain access to key information while a more strategic solution is developed,
 - a. Exploring options for e-only digital media which has a demonstrable significant demand to our readers
 - b. Discussing with publishers interim arrangements whilst ingest solutions are developed, e.g. Legal Deposit access to e-only content held by the publisher
 - c. Discussing with publishers the need for designated archival arrangements for their content in the event of the publisher going out of business or their content being withdrawn

- d. Discussing with key stakeholders in the public and private sectors the challenge of non-traditional publications, such as apps and video games, with a view to securing the long-term preservation of and access to of a selection of these works within the UK, though not necessarily within the Legal Deposit Libraries themselves
5. *Digital sound recordings.* We will review how we ensure comprehensive coverage of UK recorded sound. Previously, UK recorded sound had been excluded from the Legal Deposit Libraries Act 2003 and the 2013 Regulations because evidence at the time suggested that voluntary schemes for the deposit of this material were working sufficiently well. However, there has been a dramatic falling off of coverage via this route because of changing practices in a digital context, for example, the proliferation of individuals and small labels producing their work digitally and who do not belong to any professional association or scheme. There is now a danger that a large percentage of UK recorded sound will be lost if left to existing voluntary arrangements. The Libraries will therefore explore options for remedying this situation, and if analysis suggests it is the best solution, will advocate the extension of the Regulations to include recorded sound.

Doc 1.0 Feedback from Ian Cooke, Head, Contemporary British Publications, BL, June 2015. Doc 0.9 Feedback from CDAS representatives of Oxford, NLS, NLW, and TCD, May 2015. Doc 0.6 Feedback from Debbie Cox, Lead Curator, Contemporary British Publications, BL, April 2015; Presentation at CDAS collection development workshop, March 2015; Feedback from Sally Halper, Head, Service and Content Development, Humanities and Social Science, 27th Feb 2015; Doc 0.4 Feedback from Helen Hockx-Yu, Head, Web Archiving; Doc 0.3 Feedback from Richard Ranft, Head, Sound and Vision, 5th Feb 2015; Doc 0.2 Feedback from Alasdair Ball, Head, Collection Management; Ian Cooke, Head, Contemporary British Published Collections, 2nd Feb 2015; Doc 0.1 Richard Price, Head, Contemporary British Collections, 21st Jan 2015