**Type of Contribution: PAPER**

**Emerging Formats: understanding the impact of new publishing technologies on libraries’ collecting practices**

**Keywords: emerging formats, digital preservation, non-print legal deposit, eBook mobile apps, interactive narratives**

**Introduction**

How do you respond to innovation? How do you ensure that today’s new formats can be preserved as tomorrow’s history? In a constantly changing digital environment, cultural organisations worldwide are tasked with these very questions.

Now more than ever there are new technologies, platforms, tools, funding models and formats for creating digital publications and shaping how audiences consider a book (Mod 2018). In order to keep their offer relevant and successfully reflect the breadth of today’s digital offer, libraries must react in a timely manner, and find a way to adjust their current policies to accommodate new formats and emerging media in their collection.

In order to address these issues the British Library, together with the other five UK Legal Deposit Libraries (LDLs) set up the ‘Emerging Formats Project’ (The British Library 2018). This research explores the outcomes of this project, investigating how the emergence of new types of digital publications affect libraries’ collecting practices, and how these libraries can build their existing knowledge and capability to manage these new complex digital objects.

**Background**

This research is informed by the *Legal Deposit Libraries (Non-Print Works) Regulations 2013* (Parliament of the United Kingdom 2013)*,* which extended legal deposit to include non-print publications. The main objective of these Regulations was to help ensure comprehensive collection of UK publications, as well as safeguard them against the risk of a digital black hole, which is an especially pressing issue for born-digital publications with no print counterpart. Since 2013, the UK LDLs have been collecting a variety of born-digital material, mainly comprised of eBook, eJournals and archived UK websites (The British Library n.d)

EBooks and eJournals (and even some HTML webpages) have similar structures to what we understand as a book, but what about other publications that might not be immediately recognisable as book-like? The UK LDLs are now looking at formats in scope under Non-Print Legal Deposit (NPLD) but whose structure, content and features are more complex compared to digital publications already in their collection. These are called Emerging Formats.

**Research questions**

* How are new formats and their technical dependencies shaping libraries’ collection policies?
* How can GLAM institutions ensure they’ll be able to not only collect, but also preserve current digital outputs in a meaningful way, in order to be able to answer to user needs in the future?
* How do we determine user access needs for content that has yet to be collected?

**Emerging Formats** **Project**

For the purpose of the project emerging formats were defined as born-digital publications with no print counterpart. They have strong software and hardware dependencies and often consist of more than one media type. They are comprised of non-standard formats that might never standardise. The environment in which they are created is one of continual change, and most of these new formats are already at risk of rapid obsolescence (Smith and Cooke 2017).

The UK LDLs chose to focus on three emerging formats in particular:

* eBook mobile apps: digital books published as mobile apps. They have strong hardware and software dependencies and often make use of interactive features characteristic of mobile technology;
* web-based interactive narratives: online text-based stories which require the reader’s active input to determine how the narrative unfolds;
* structured data: publications that include both the data and the interfaces necessary to organise and access the data.

The main goal of the emerging formats project was to devise a system to identify, collect, describe, preserve, make discoverable and available complex objects within scope of NPLD in a timely manner. This was to be achieved by means of different resources and collaborations (with creators, users, other memory organisations, etc.) in order to provide access in an authentic and meaningful way and preserve content over the long term.

This paper explores the challenges linked to each of these aspects, including the methodology used and key findings.

*Identify*

One of the first challenges in the collection management workflow is that of identifying new publications and tools and determining what’s in scope for collection. Emerging formats proved to be especially challenging as there are no comprehensive directories, and when there are lists available, these tend to be global and not UK-specific. The project began with a scoping phase to research the market for publications the libraries weren’t already collecting, and with building a sample collections of works made with in scope formats. Already-established workflows for other digital objects were considered, to test possible overlaps and assess gaps in knowledge and tools (Day et al. 2018). Key sources of information were identified in constant engagement with creators and publishers, as well as an active participation in relevant new media communities and sharing expertise with other cultural organisations dealing with similar challenges.

*Collect*

Libraries must collect versions that support preservation and access over the long term. Challenges arise when more than one version is available for the same publication: How do we determine what the best version to collect is?

During the project different methods to collect content under NPLD were identified, with some methods more suited to certain type of content. Sometimes a “mixed approach” – collecting more than one version to ensure preservation and access, or using more than one collection method to capture different parts of a publication – might be the best solution (Clark 2019). Software and hardware dependencies, as well as issues related to Digital Rights Management, had to be considered not just for the collection stage, but throughout the whole management process.

*Preserve*

Libraries need to preserve content over the long term, to enable access and to maintain the integrity and authenticity of the work. Collection profiles and preservation assessments were conducted for all prioritised formats (Pennock, forthcoming).

Many of these file formats are not represented in registries and are unsupported by validation tools, which make ensuring authenticity and integrity more challenging. Different methods for recording authorial intent were taken into consideration (including preserving contextual information around a publication) in order to maximise the chances of preserving authentic copies. It was also noted that access to preserved files might require additional software and the development of emulators.

*Describe and Discover*

The emerging formats collection within the library needs to be easily discoverable and to offer readers enough information to support use. Rich descriptive metadata (e.g. technical requirements, operating systems, hardware information, and interactive methods) is necessary to support future access, provide a record of any loss of functionality and facilitate discovery between contextual information and the publication itself.

*Access*

Research in the field of new digital publications is also emerging, which makes anticipating user needs especially difficult. The British Library conducted a targeted UX testing and focus groups to understand user expectations for access to emerging formats. The UX research identified a strong interest in the added value of a curated collection of emerging formats, as well as support for the collection of contextual information around publications.

These findings helped the LDLs shape the collection management methodology for emerging formats, including what and how they collect complex digital objects. They will also be used to inform access requirements and ensure that these publications are easily accessible and discoverable, in order to provide a relevant and meaningful user experience.

**Conclusions**

The research findings support the value of libraries collecting and preserving complex born-digital publications to represent the changing nature and cultural diversity of the UK digital landscape. Collection policies must take into consideration technical dependencies and access requirements, as well as anticipate research needs in a digital environment. In a rapidly evolving landscape, libraries need to adjust the way they deal with born-digital material, in order to maximise its potential and avoid limiting it within the boundaries of unsuitable print access policies (Gooding et al. 2019).

**REFERENCES**

Clark, Lynda. 2019. “Collecting Interactive Fiction.”*UK Web Archive blog, The British Library*, March 29, 2019. <https://blogs.bl.uk/webarchive/2019/03/archiving-interactive-fiction.html>.

Day, Michael, Pennock, Maureen, Smith, Caylin, Jenkins, Jeremy, and Cooke, Ian. 2018. “Preservation Planning for Emerging Formats at the British Library.” iPres 2018 conference paper. <https://doi.org/10.5281/zenodo.1303008>.

Gooding, Paul, Terras, Melissa and Berube, Linda. 2019. “Towards User-Centric Evaluation of UK Non-Print Legal Deposit: A Digital Library Futures White Paper.” <http://elegaldeposit.org/dlf-white-paper>.

Mod, Craig. 2018. “The ‘Future Book’ Is Here, but It’s Not What We Expected”. *WIRED*, December 20, 2018. <https://www.wired.com/story/future-book-is-here-but-not-what-we-expected/>.

Parliament of the United Kingdom. 2013. “The Legal Deposit Libraries (Non-Print Works) Regulations 2013.” Accessed 29 September 2019. <https://www.legislation.gov.uk/ukdsi/2013/9780111533703/contents>.

Pennock, Maureen, May, Peter and Day, Michael. forthcoming. “Considerations on the acquisitions and preservation of eBook mobile apps.” iPres 2019 conference paper.

Smith, Caylin, and Cooke, Ian. 2018. “Emerging Formats: Complex digital media and its impact on the UK Legal Deposit Libraries.” *Alexandria: The Journal of National and International Library and Information Issues* 27(3): 175–187. [https://doi.org/10.1177/0955749018775878](https://doi.org/10.1177%2F0955749018775878).

The British Library. 2018. “Projects | Emerging Formats.” Accessed 29 September 2019. <https://www.bl.uk/projects/emerging-formats>.

The British Library. n.d. “Legal Deposit.” Accessed 29 September 2019. <https://www.bl.uk/legal-deposit>.