Emerging Formats

Ian Cooke, Head of Contemporary British Publications

Caylin Smith, Legal Deposit Libraries Senior Project Manager

#PASIG17 September 11, 2017

@IanCooke13, @caylinssmith

BRITISH
LIBRARY

Legal Deposit Libraries Act 2003

- Responding to "digital black hole"
- Established principle that Legal Deposit **could** extend to 'works published in media other than print'
- Addressed copyright restrictions on collecting digital content
- Excluded sound recordings and film
- Extended protections of Legal Deposit to non-print works

UK Non-Print Legal Deposit (NPLD)

- NPLD Regulations came into effect in 2013

- NPLD allows the Legal Deposit Libraries to collect, preserve, and provide access to content in digital form that they would previously acquired in print.

- This content is currently comprised of:

  - eBooks

  - eJournals

  - Websites

  - Sheet music

  - Map datasets

Deposit with British Library (Boston Spa for NPLD content)

Content then shared and distributed to four storage nodes:
Boston Spa (BL), St Pancras (BL), Aberystwyth (NLW), and Edinburgh (NLS)

Access at each node, and by secure network for:
Cambridge University Library, Trinity College Dublin, and the Bodleian Libraries

Ingest, Metadata, Access and Printing

- We can ingest and preserve eJournals and eBooks from a variety of sources
- We can ingest and preserve websites
- Metadata can be harvested from the BL by all Legal Deposit Libraries
- Each LDL now has the ability to access eJournals, eBooks, and the Web Archive
- Printing is available across the LDLs dependent upon local infrastructure

Background on Emerging Formats

'Emerging formats' refers to media and formats for which collection management is not currently supported either by existing capabilities or through planned capability development.

This project is situated within a British Library and LDL context as well as other projects, including

- The BL's IT transformation

- Voluntary deposit of UK sound recordings

- The review of NPLD regulations

- Other NPLD projects (e.g. Digital Maps; Digital Sheet Music)

Emerging Formats Content - Prioritisation

Does the format have identifiable characteristics – can we define it?

Evidence of reader demand

Creators are willing to engage with us

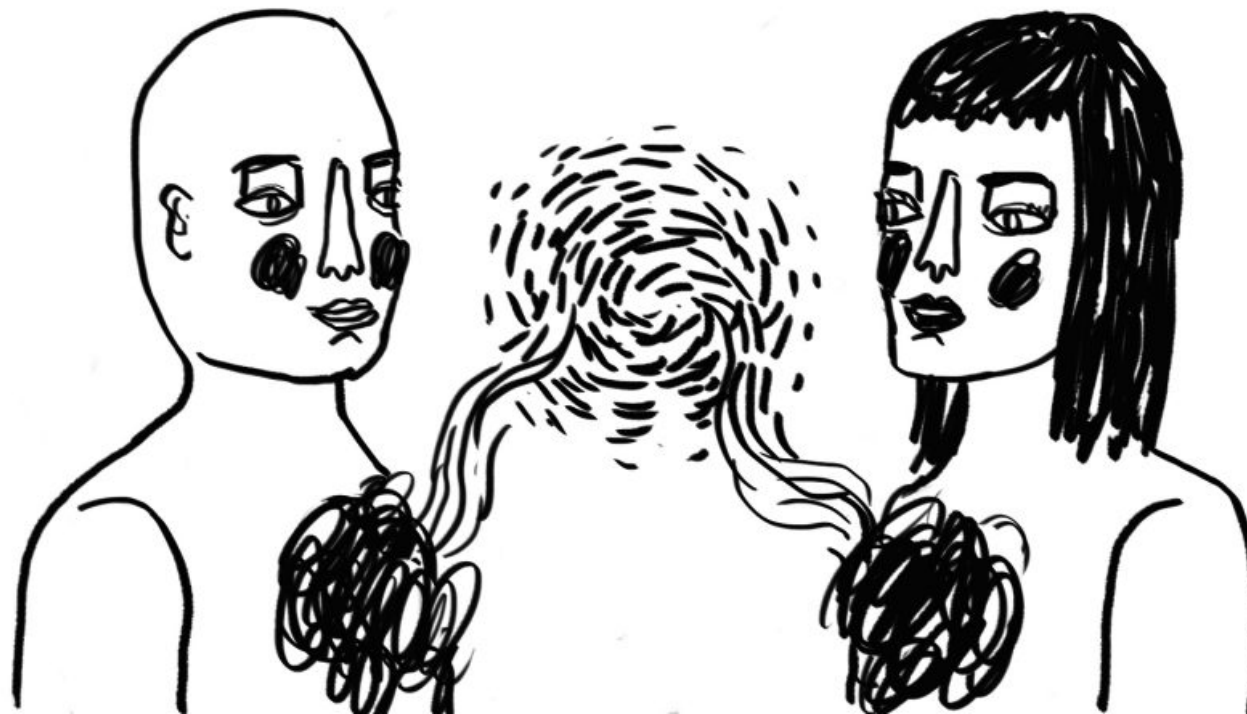Risk of loss identified

Emerging Formats Content

- Books as apps
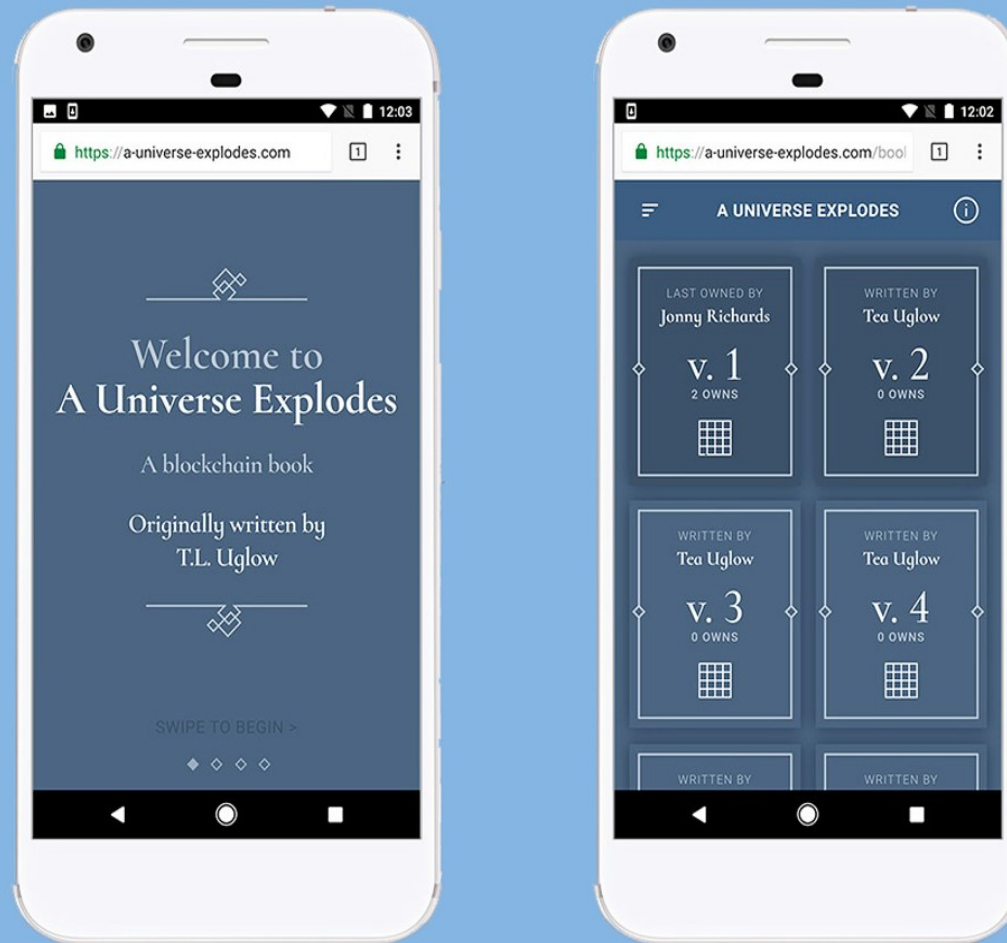- Interactive narratives
- Structured data

Books as Apps

- In the UK, the total amount published is probably in thousands (numbers hard to find out)
- Applies to both academic and fictional content, especially children's publishing
- There's crossover into gaming / book apps being marketed and described as games
- Published on iOS and/or Android, plus other app formats
- Looking into the impact of EPUB3 file format

Goldilocks and Little Bear by Nosy Crow

The Group Therapy Program by OWN IT!

A Universe Explodes by TL Uglow
Part of Visual Editions' Editions at Play

Interactive Narratives

- To be considered more as a genre than a type of format
- Looking into whether web harvesting tools can be used to capture content

The Wondering Lands of Alice by Off Our Rockers

Structured Data

- Smaller in numbers compared to apps

- Interested into linked data from journal articles acquired

- Interested in data-driven publications (e.g. Lloyd's List)

UK Parliament Gazette, https://www.thegazette.co.uk/

Challenges Identified

- What's out there?

- Territory in a networked world

- Developing an end-to-end scalable solution

- Lack of standards for how content and metadata are produced

- Increased dependence on software and hardware for preservation and access

- Technical knowledge of how content was created; how it can be re-built

- Identifying whose content we need to collect

- Defining what the Library will actually acquire

What's Next?

- Collect sample content
- Create methodology for carrying out analysis
- Make networks (e.g. Autumn 2017 research event)
- Determine requirements collection management
- Write Emerging Formats Strategy

Thank you!

Ian Cooke (ian.cooke@bl.uk)
Caylin Smith (caylin.smith@bl.uk)