E-book Preservation

Barbara Sierman, DigitalPreservation.nl

Springer Webinar 10-03-2021
Digital Preservation

Preservation of knowledge: about the past and for the future

Goals summarized in the ISO 14721 The Open Archival Information System (2005; 2012; 2021)

Generic reference model: Standard internationally widely adopted in a variety of domains

Translations in many languages Spanish, French, Chinese etc.
Some key concepts

Appropriate information from the information Producers

Designated community: the audience intended

Information should be independently understandable

Archive will obtain sufficient control for preservation

Authenticity, supported by provenance / fixity information

Presented object is traceable to original object

The rights to do preservation and to give access
Preservation Environment

Started with Space Agencies, Libraries and Archives around 1990

Since a few years awareness for preservation in research data management

2016

2020

Findable, Accessible, Interoperable, Reusable

TRUST

Transparency, Responsibility, User-focus, Sustainability, Technology

b.sierman@digitalpreservation.nl
E-book preservation

By whom?

Libraries (National, University etc.)
Publishers
Specific archives like Portico and CLOCKSS

Changes in the concept “publication” might lead to other custodians
What is an e-book?

Disclaimer: Concept “publication” is less firmly defined nowadays

Digitized or born-digital

Variety of formats, no longer equivalent to pdf

EPUB standard, preservation objectives included open format but ISO standard behind paywall

Propriety formats (Kindle, KOBO, Apple iBooks etc.)
What do we need to preserve?

Identification and validation of the file format: EPUBchecker, veraPDF etc.

(Enough? What about Flash or Javascript in EPUBs?)

Open file format, propriety formats are difficult to preserve

Check for security measures in the e-book

Submission agreement with information about preservation rights

Context information

Metadata
What do we need to give access

- Proof of authenticity
- Satisfy the designated community / communities
- Faithful representation via readers
- Adapt to reader habits: preferences for EPUB instead of PDF
Some challenges?

• Traditional concept of “book” or “publication no longer firmly defined

• File format developments
  • EPUB format accepted as preservation format (Library of Congress) but still some doubts
  • Preservation community monitors developments

• What to preserve?
  • One format, many representations (“adaptable publications”)
  • What about “original look and feel”
  • External links (Link rot, Reference rot)
  • Context: original publisher environment
To summarize

E-book preservation requires:

• Ambition and awareness

• Money

• Preservation Watch: monitoring the developments constantly

Thank you!