Born digital news collections

Other libraries approaches

Elisa Villanueva, Koninklijke Bibliotheek
- Part of this news content is already lost or in its way to disappear (not archived, constant changing, broken links).
- Important part of our cultural heritage.
- Survey students / staff Missouri school of Journalism.
Born Digital

- Originated as a digital product since its beginnings.
- Born digital content is distinct from digital content, which is created through the digitization of analog content.
- Different lifecycle
- Library’s need a different model to preserve news-content.

- Repetitive, redundant. It is vital to understand thoroughly with which type of content we are dealing.
- Different production system, different approach, this new model will consider new and more decisions when developing the collection.
SAME base: Editorial system. MAIN DIFFERENCE: Diversity of outputs / Constant evolution
Bigger scope for the selection criteria. More choices, more complex. Example: Which output, version, frequency??
### SELECTION

- **Format:** What should we prioritize?
- **Layout v/s content? Quantity v/s quality?**
- **Frequency:** Capture
- **Accuracy:** According to the original
- **Completeness:** Intended scope

### TECHNICAL

- **Outputs**
- **Collecting procedure**
- **Metadata**
- **Paywalls**
- **Storage**
- **Obsolescence**

### RELATIONSHIP WITH PUBLISHERS

- **Copyrights**
- **Upload system**
- **Access**
- **Agreements**
- **Workflow**

---

**SELECTION**

Format: What should we prioritize?

Layout v/s content? Quantity v/s quality?

Frequency: Capture

Accuracy: According to the original

Completeness: intended scope

**TECHNICAL**

Collecting procedures and preservation technologies.

Level of metadata (search).

Technical obsolescence (future access).

**RELATIONSHIP WITH PUBLISHERS**

Because of the different outputs and their different flows, there would be necessary to set a new type of agreements with publishers.

Also, every publisher has a different procedure: difficulties to set a standard model.
Research

What is the approach other libraries or cultural heritage institutions took (or are taking) during the creation of their born digital news collection?

• Report which illustrates which could be the best practices, main challenges, issues, etc. per institution consulted.

• Main purpose: Use this as a starting point in order to think in standards for a possible model in the KB.

SO... these aspects would work as a base for developing our research approach.
## Methodology

Institutions consulted:

<table>
<thead>
<tr>
<th>National Library of Denmark</th>
<th>National Library of Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swiss National Library</td>
<td>Biblioteque nationale de France</td>
</tr>
<tr>
<td>National Library of South Africa</td>
<td>National Library of Germany</td>
</tr>
<tr>
<td>National Library of Luxembourg</td>
<td>University Library Kentucky</td>
</tr>
<tr>
<td>British Library</td>
<td>University Library of Texas</td>
</tr>
<tr>
<td>Library of Congress</td>
<td></td>
</tr>
<tr>
<td>National Library of Australia</td>
<td></td>
</tr>
<tr>
<td>National Library of Croatia</td>
<td></td>
</tr>
</tbody>
</table>
The results showed us that these are the main approaches for BDNC.

WEB ARCHIVING

Completeness: France web and print.

ISSN British Library

Paywalls: Croatia do not collect web sites behind pay walls because of the technical issues.

Flash, videos: BL and Bnf.

Accuracy: Impossible to reproduce web sites exactly.

PDF

Germany experience

OTHERS

Sweden
Best practices

• Collecting databases (e.g.: articles rather than entire web sites).
• PDFs (access, longterm preservation).
• RSS feeds (constant updates).
• Closer relations with publishers (dialogue, direct transfer of content).
• Third party mediator (intermediator, distributor).
Conclusions

• Importance of partnership with publishers (workflow/standards).
• Understanding that collecting born digital has to be through a different model (new formats, new lifecycle, new actors). Libraries have to make more and complex decisions in order to think in new policies.
Questions?

Comments?

Thanks!