Experimenting with ‘Collections as Data’ at KBR: an interdisciplinary collaboration

Sally Chambers and Frédéric Lemmers
What is ‘Collections as Data’?

Collections as Data

Providing data-level access to digital collections is a primary challenge for undertaking digital humanities research. In the United States, the flagship initiatives, ‘Always Already Computational: Collections as Data’ and ‘Collections as Data: Part to Whole’, define ‘Collections as Data’ as a “conceptual orientation to collections that renders them as ordered information, stored digitally, so that they are inherently amenable to computation”. The initiative was established to document, exchange experience and share knowledge to encourage cultural heritage institutions to implement ‘collections as data’ in their own institutions. DATA-KBR-BE will kick-start the implementation of ‘Collections As Data’ in Belgium.

https://collectionsasdata.github.io
Data-level access

Providing access to the underlying files of digitised cultural heritage resources to facilitate data analysis by means of tools and methods developed in the field of digital humanities.

```
01
  alto
    KB_JB840_1919-04-01_01-00001.xml
    KB_JB840_1919-04-01_01-00002.xml
  jpg
    KB_JB840_1919-04-01_01-mets.xml
    KB_JB840_1919-04-01_01.pdf
  pdf
    KB_JB840_1919-04-01_01-00001.pdf
    KB_JB840_1919-04-01_01-00002.pdf
  tif
    KB_JB840_1919-04-01_01-00001.tif
    KB_JB840_1919-04-01_01-00002.tif
```
Introducing the DATA-KBR-BE project

Facilitating data-level access to KBR’s digitised and born-digital collections for digital humanities research

From Collections to Data
Interdisciplinary Research Scenarios

- **Collective Action Belgium** led by GhentCDH, focuses on social history in the Interbellum and World War Two period and aims to trace the dynamics of contention, strikes, demonstrations and other forms of collective action in Belgium as reported in Belgian newspapers;
- **The feuilleton in Belgium**, led by ACDC, focuses on literary studies in the period 1830–1930 and aims to map the publication of literature in Belgian newspapers across the first century of the Belgian nation state;
- **History of Belgian Journalism**, led by ULB and KBR, focuses on media history from 1886 until now and aims to trace the history of Belgian journalism through the lens of critical discourses about journalism as in Belgian newspapers.

1) Collective Action Belgium

**Discipline:** Social history

**Period:** interbellum & WWI

**Topics:** tracing the dynamics of contention, strikes, demonstrations and other forms of collective action in Belgium
2) Feuilleton in Belgium

**Discipline:** Literature  
**Period:** 1830 - 1930  
**Topics:** mapping the publication of literature in Belgian newspapers across the first century of the Belgian nation state
3) History of Journalism in Belgium

**Discipline:** Media History

**Period:** 1830-1970

**Topics:**
1) a database of Belgian journalists
2) the history of Belgian journalism through the lens of critical discourses in Belgian newspapers
Towards a ‘Collections as Data’ workflow

1) Collective Action Belgium

**Discipline:** Social history

**Period:** interbellum & WWI

**Topics:** tracing the dynamics of contention, strikes, demonstrations and other forms of collective action in Belgium
## General Strikes in Belgium

<table>
<thead>
<tr>
<th>Year</th>
<th>Date</th>
<th>Title</th>
<th>Cause</th>
<th>Number of strikers</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1893</td>
<td>12–18 April</td>
<td>Belgian general strike of 1893</td>
<td>Franchise reform</td>
<td>200,000</td>
<td>Successfully led to the establishment of universal male suffrage with plural votes. Thirteen strikers were killed and socialist leaders were briefly arrested.</td>
</tr>
<tr>
<td>1902</td>
<td>10–20 April</td>
<td>Belgian general strike of 1902</td>
<td>Franchise reform and an end to plural vote</td>
<td>350,000</td>
<td>Failed to achieve the abolition of the plural vote as Catholics and Liberals united to oppose constitutional reform. The Belgian Workers' Party had been reluctant to support the strike and it soon descended into violence in Brussels and parts of Wallonia. 12 workers and one policeman were killed. Union membership dropped sharply in its aftermath.</td>
</tr>
<tr>
<td>1913</td>
<td>14–24 April</td>
<td>Belgian general strike of 1913</td>
<td>Franchise reform</td>
<td>400,000</td>
<td>Carefully planned to avoid the same problems as 1902, the strike gained the promise of electoral reform but its proposals were postponed by the outbreak of World War I and the subsequent German occupation. The policy was finally adopted in 1919.</td>
</tr>
<tr>
<td>1932</td>
<td>7 July–9 September</td>
<td>Belgian general strike of 1932</td>
<td>Pay, working hours and unemployment insurance</td>
<td></td>
<td>Began after a spontaneous strike by coal miners in the Borinage and involved Communist agitation following a severe decrease in living standards and real wages during the Great Depression. Two people were killed during the strike.</td>
</tr>
<tr>
<td>1936</td>
<td>2 June–2 July</td>
<td>Belgian general strike of 1936</td>
<td>Working hours, paid holiday, union reforms</td>
<td>500,000</td>
<td>Broke out at the port of Antwerp and led to the creation of a National Labour Conference. Although influenced by the French Popular Front and held against the backdrop of the Spanish Civil War, it was also supported by Catholic trade unions.</td>
</tr>
<tr>
<td>1950</td>
<td>24 July–3 August</td>
<td>Belgian general strike of 1950</td>
<td>&quot;Royal Question&quot;</td>
<td>700,000</td>
<td>Chiefly active in Wallonia, the strike contributed to the abdication of King Leopold III on 1 August 1950. At least four strikers were killed.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Title</th>
<th>Journal No.</th>
<th>Digitally availability</th>
<th>1886</th>
<th>1893</th>
<th>1902</th>
<th>1913</th>
<th>1932</th>
<th>1936</th>
<th>1950</th>
<th>Extraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vooruit: socialistisch dagblad</td>
<td>JB 809</td>
<td>1884-1889; 1901-1902; 1911-1950</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Het Volk : antisocialistisch dagblad</td>
<td>JB 785</td>
<td>1911-1916; 18 apr. 1918-31 maart 1927; 1931-1950.</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Vaderland</td>
<td>JB 310</td>
<td>16 maart 1910-30 dec. 1913.</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Le Peuple: organe quotidien de la démocratie socialiste</td>
<td>JB837</td>
<td>1885-1907; 1911-1914; 1918-1940; 1944-1950.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Le Vingtième Siècle</td>
<td>JB729</td>
<td>6 juin 1895-13 mai 1940</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>La Meuse : journal de Liège et de la Province</td>
<td>JB638</td>
<td>1856; 1858-1882; 1884-5 août 1914; 28 nov. 1918-11 mai 1940; 9 sept. 1944-1950.</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
### Importance of historical context

<table>
<thead>
<tr>
<th>Titel</th>
<th>Vooruit : socialistisch dagblad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bibliografisch adres</td>
<td>Gent, 1884-1992</td>
</tr>
<tr>
<td>Materiële beschrijvi...</td>
<td>f°</td>
</tr>
<tr>
<td>Periodiciteit</td>
<td>Dagblad</td>
</tr>
<tr>
<td>Belgicapress</td>
<td></td>
</tr>
<tr>
<td>Bezit</td>
<td></td>
</tr>
</tbody>
</table>

“The newspaper was founded in Ghent in 1884 with the support of the socialist cooperative Vooruit. When the Parti ouvrier belge - Belgische Werkliedenpartij ("Belgian Workers Party", POB-BWP, 1885) was founded, it was recognised as its official organ for the Flemish part of the country. The Ghent socialist leader Edward Anseele was editor-in-chief, but he also worked as a typographer. Vooruit was published under German censorship during the two world wars. Having reached its peak in the 1950s, it began a slow decline. It was succeeded by the daily De Morgen in 1978.”

https://opac.kbr.be/LIBRARY/doc/SYRACUSE/15981364
Data Science

Document Layout Analysis using Belgica Press

Dilawar Ali & Steven Verstockt, IDLab, UGent
Dilawar Ali, Kenzo Milleville, Alec Van de Broeck & Steven Verstockt, IDLab, UGent
Corpus Building with Artificial Intelligence

Feuilleton in Belgium

**Discipline:** Literature  
**Period:** 1830 - 1930  
**Topics:** mapping the publication of literature in Belgian newspapers across the first century of the Belgian nation state

Wout Dillen, University of Antwerp/Borås
Feuilleton in Belgium

Feuilletons: a combination of distinct visual and textual features

- Regularly appeared in the bottom third of the page
- Often extended the full breadth of the page and appeared on the front or the second page
- Distinctive titles, ended with words such as ‘to be continued’ (Wordt voortgezet), Number of the serial part was often mentioned (e.g. 5e vervolg)

Next Steps: To train a model to identify and extract the Feuilletons into a corpus.
Data Science

NewspAlper

Metadata extraction of historical newspaper archives.

https://tw06v072.ugent.be/kbr/

Dilawar Ali, Kenzo Milleville, Alec Van de Broeck & Steven Verstockt, IDLab, UGent
Oberbichler, S. et al. (2021) *Integrated interdisciplinary workflows for research on historical newspapers: Perspectives from humanities scholars, computer scientists, and librarians.* JASIST, August 2021. [https://doi.org/10.1002/asi.24565](https://doi.org/10.1002/asi.24565)
Thank you!

DATA-KBR-BE


- Frédéric Lemmers (Frederic.Lemmers@kbr.be)
- Sally Chambers (Sally.Chambers@kbr.be)