



europaena  
newspapers



---

# Europeana Newspapers Project and Beyond

Berlin, January 31, 2014

Hans-Jörg Lieder, Project Coordinator - Berlin State Library

Staatsbibliothek zu Berlin

# Overview

1. Why newspapers
2. Core mission ENP
3. Presentation on the European level
4. Services
5. The 20th century

# Why newspapers? (and why not yet)

- Not meant to be preserved
  - Stubborn, if not wicked material (paper, ink, fonts, layout, library binding, wrongly sorted or missing editions, ...)
  - Sheer quantity is intimidating
- 
- Reflects immediacy and fast pace of public discourse
  - Covering (almost) every aspect of the world (global, regional, local news, culture, sports, science, politics, economy, advertising, obituaries, ...)
  - Vast cross-domain interest



# Europeana Newspapers Project Partners

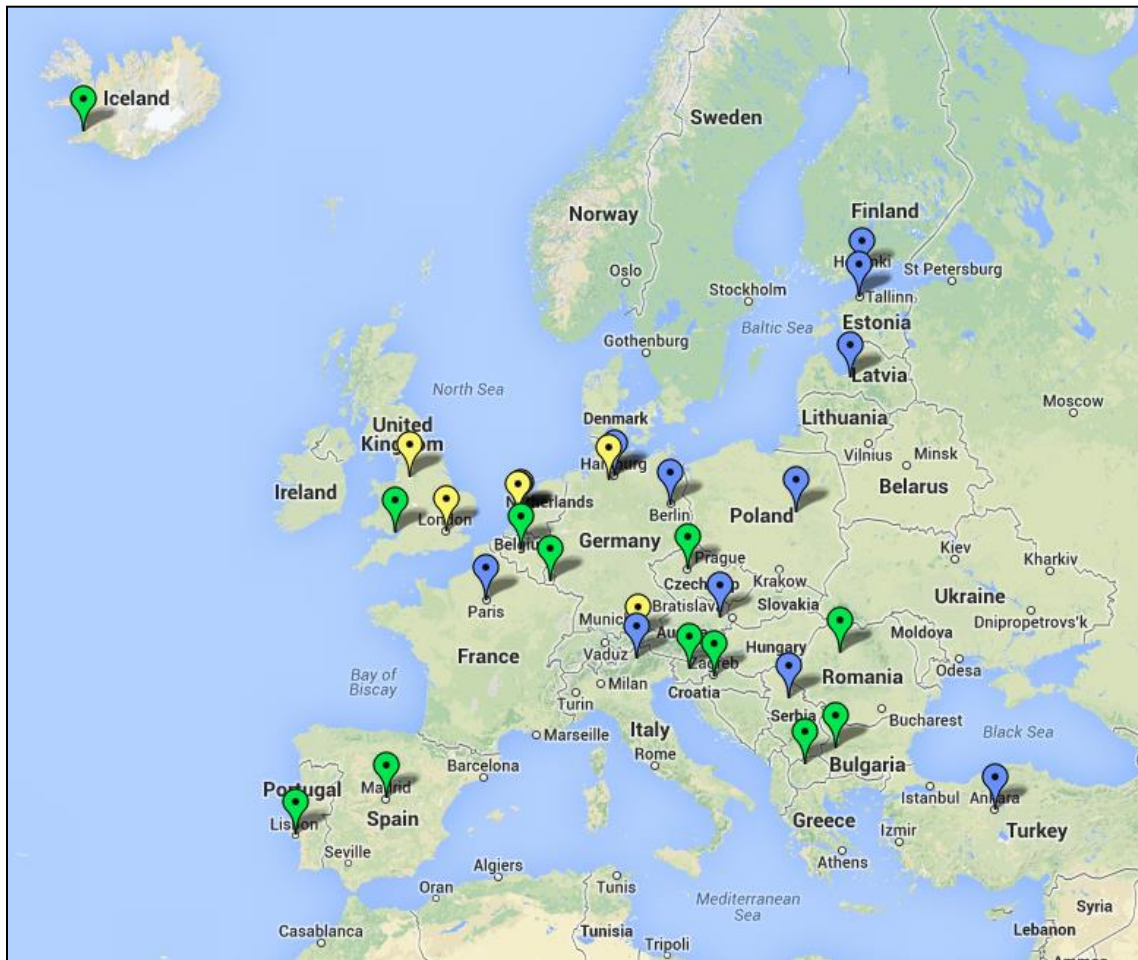
01. State Library Berlin, Germany
02. National Library of the Netherlands
03. National Library of Estonia
04. National Library of Austria
05. National Library of Finland
06. State and University Library Hamburg, Germany
07. National Library of France
08. National Library of Poland
09. University of Salford
10. CCS Content Conversion Specialists GmbH
11. Stichting LIBER, Netherlands
12. National Library of Latvia
13. National Library of Turkey
14. University Library of Belgrade
15. University of Innsbruck
16. State Library Dr. Friedrich Tessmann, Italy
17. The British Library, UK
18. Europeana Foundation, Netherlands



This project is partially funded under the ICT Policy Support Programme (ICT PSP) as part of the Competitiveness and Innovation Framework Programme by the European Community [http://ec.europa.eu/ict\\_psp](http://ec.europa.eu/ict_psp)



# Who



Blue– Providing Content

Yellow –Providing Technical Services

Green – Associate Partners

# Content

Country	Partner	Titel	Startdate (overall)	Enddate (overall)
France	BnF	80	1814	1944
Germany	SBB	6	1872	1940
	SUB-HH	16	1721	1945
Netherlands	KB	203	1618	1900
Italy	LFT	15	1813	1949
Estonia	NLE	43	1852	1944
Finland	NLF	11	1900	1910
Latvia	NLL	117	1868	1955
Poland	NLP	118	1914	1939
Turkey	NLT	22	1818	1928
Austria	ONB	275	1686	1945
Serbia	UB	45	1830	1944

**20 languages**

**ca. 950 titles**

**ca. 10m pages refined**

**8m OCR**

**2m OLR**

**2m NER**

**vast amount of newspapers from 1914-18**

# Further to content

1. Defining metadata standards (images, texts, structural metadata, enrichments)
2. Developing quality evaluation and prediction tools to estimate OCR and OLR results
3. Defining aggregation and delivery mechanisms to transfer data from local providers to TEL and Europeana

More at German Information Day: February 27/28, Berlin

# The European Newspapers Browser

## First Iteration:

- Basic text search
- Filtering of results by date, country, newspaper, language, library
- OCR results shown
- Zoom able version of full image
- Clickable links between full text and image (*where possible*)
- Link to newspaper source library

## Second Iteration

- Added information about particular titles
- See what was published on a particular day
- Search over titles (not just text)
- Other browse-able visualisations of publication and library source
- Search / browse via entities



# The European Newspapers Browser

The screenshot displays the website's navigation and search features. At the top, there are links for 'About', 'Membership', 'For Current Partners', 'newspaper', 'Log out', and 'English (en)'. The main header includes 'The European Library' logo and a search bar with the text 'Search all European Library content...'. Below this, there are tabs for 'HOME', 'DISCOVER', and 'SERVICES FOR RESEARCHERS'. The main content area is titled 'Newspapers BETA' and includes a search bar for 'Search Newspapers' and a secondary search bar for 'Search within historic newspapers...'. There are also buttons for 'Newspapers Home', 'Content', and 'Project'. A section titled 'Browse by Issue Date' features a calendar grid with a 'Select year' dropdown. Another section, 'Explore Newspapers', shows a collage of newspaper front pages. A third section, 'Browse by Newspaper Title', lists titles such as 'Wiener Zeitung (29,018 issues)', 'Bote für Tirol (24,088 issues)', 'Bozner Zeitung (18,420 issues)', and 'Morgenzeitung (9854 issues)'. A fourth section, 'Browse by Country', shows a map of Europe with a green highlight on Germany.



This project is partially funded under the ICT Policy Support Programme (ICT PSP) as part of the Competitiveness and Innovation Framework Programme by the European Community [http://ec.europa.eu/ict\\_psp](http://ec.europa.eu/ict_psp)



# The European Newspapers Browser

Everything
Newspapers
Collections
Full text
Remote Search

REFINE

You are looking at **10** items out of **280,113**

[Wiener Sonntags-Zeitung / 22 Nov 1868/ Page 1](#)


!U> ««,» . «. »»f »|» ; «««KW »|» »»rL» »>|«tt. ««|iU««« » . »»>««««««««« |» . «. ^ l.xmi : Str»»okx»s««, (vir-jl-vi, Ssr VSr»s.l »»«««« » lrr. «. V,

[Wiener Zeitung / 20 Dec 1797/ Page 1](#)

Wie«. hab«» dem K.L.Rakhvüid Hofaxestea, H-n. Ferd.MAker, Fdkyhnm von'und zUMÄeZ?, WRÄ»- BM sckes altadettgn NLstammens, «nd als ømen «veiterv Beweis der durch stiae k ver?«-txa=j, Geitzrichtik«! an Tag ge, legte Thätigkeit, jdh erw«kbe=en aller- höchstni Zvfried«,heit» die Würde eines K K

[Wiener Zeitung / 31 Dec 1777/ Page 1](#)

Mittwoch, den zr. Christin«?. 1777. — — — — — iz^ ,>««/? . ^>?.s//s »K »»««/»» ^scr'Ac'!>./.' , cs«y?^r. Nr.r,k.



Search within results GO

Contributor ▼

- [Teßmann Library](#) (95,162)
- [Austrian National...](#) (91,894)
- [National Library ...](#) (83,361)
- [National Library ...](#) (9696)

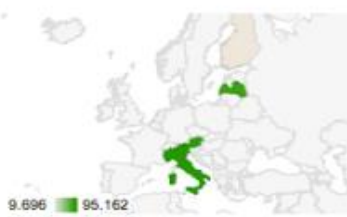
Newspaper Title ▼

Years of Publication ▼

Country of Contributor ▼

---

COUNTRY OF CONTRIBUTOR



9.696  95.162

[World](#) [Europe](#) [Western Asia](#)  
[Zoom](#)

# The European Library / Europeana

- Full texts will be available at [www.theeuropeanlibrary.org](http://www.theeuropeanlibrary.org)
- Images and metadata will be available at Europeana
- In future?

# A word about services

- Collection of texts = knowledge base
- Quality of texts and level of aggregation crucial for level of services
- Size matters – big and small
- Open services to be customized by users
- Text search, search in structured texts, data mining (frequency analysis, semantic world field analysis, searches for „networks“, geographical searches, ...)
- Librarian resources: islands of knowledge in an ocean of information
- Non-librarian resources: the world as the world sees it
- Think visual!

# Mind the gap: towards the contemporary

## Challenges:

- Legal stuff (IPR, „personal rights“)
- Business models and uncertainties

## Options:

- Seek understanding with rights holders
- Cooperate with authors associations
- Cooperate with publishers (e.g. digitization projects)



**Thank you!**

**Contact:**

**[hans-joerg.lieder@sbb.spk-berlin.de](mailto:hans-joerg.lieder@sbb.spk-berlin.de)**

**[ulrike.koelsch@europeana-newspapers.eu](mailto:ulrike.koelsch@europeana-newspapers.eu)**

For more information, please see:

[www.europeana-newspapers.eu](http://www.europeana-newspapers.eu)

Twitter (@eurnews)

Facebook: <https://www.facebook.com/EuropeanaNewspapers>



This project is partially funded under the ICT Policy Support Programme (ICT PSP) as part of the Competitiveness and Innovation Framework Programme by the European Community [http://ec.europa.eu/ict\\_psp](http://ec.europa.eu/ict_psp)

