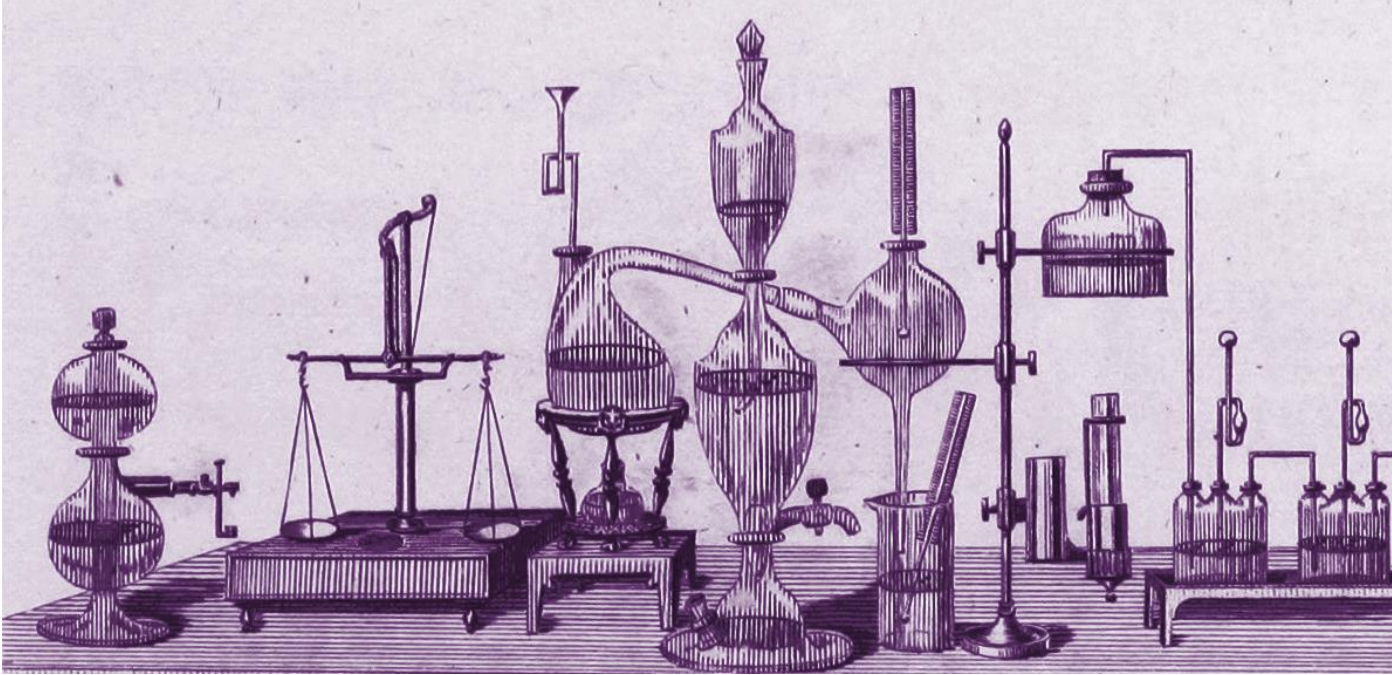


KBR 8



KBR Digital Research Lab: towards data-level access to KBR collections for research

Sophie Vandepontseele, KBR and Christophe Verbruggen & Sally Chambers, Ghent Centre for Digital Humanities

Towards the KBR Digital Research Lab

- Libraries as humanities labs?
- Towards the KBR Digital Research Lab
- Understanding user needs
- Potential research pilots
- Data-level access to collections

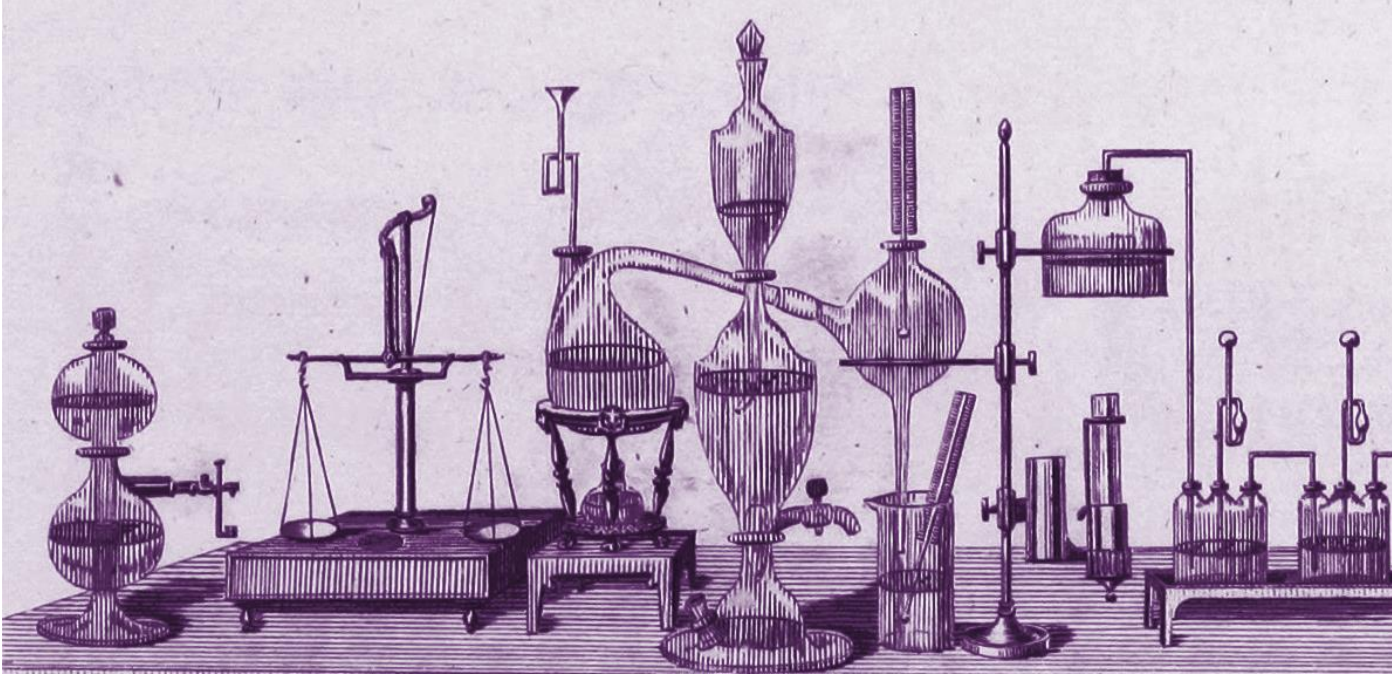
KBR 



**GHENT
UNIVERSITY**



FED-*t*WIN



KBR 8



**GHENT
UNIVERSITY**

Libraries as humanities labs?

Founded in 2013

LIBRARY
HSILIRB

British Library Labs

Experiment with our
digital collections

<http://labs.bl.uk>



@BL Labs @britishlibrary @labs@bl.uk @BL DigiSchol

British Library Labs (BL Labs)

- **Promotes, inspires, and supports the use of the British Library's digital collections and data**
- **Works on projects with researchers, developers, educators, entrepreneurs and artists from around the world**
- **Provides insight into the emerging practice of digital research and helps shape the provision of the British Library's digital services, tools, collections, and data.**
- **Ensures that the intellectual digital heritage we hold is accessible to everyone for research, inspiration and enjoyment**



Library of Congress Labs (LC Labs) Founded Sept. 2017

LIBRARY
LIBRARY
OF CONGRESS

LABS

People The Signal Blog Experiments Meta Events LC for Robots



Welcome to Labs!

A place to encourage innovation with Library of Congress digital collections.

Gallinger, M. & Chudnov, D. (2016)
Library of Congress Lab. Library of Congress Digital Scholars
Lab Pilot Project Report

https://labs.loc.gov/static/portals/labs/meta/images/DChudnov-MGallinger_LCLabReport.pdf

<https://labs.loc.gov/>

LIBRARY OF CONGRESS LAB

*Library of Congress Digital Scholars Lab Pilot Project
Report*


The Lab envisioned by the authors in this report is based on successful models that support access and use of digital collections, which have been drawn from interviews with experts and scholars and shaped by practical experience. The recommendations made in this report consider the Lab as part of the unique structure and offerings of the Library of Congress.


Michelle Gallinger and Daniel Chudnov


12/21/16

Join us and explore the KB's digital treasure trove

The KB Lab hosts all experimental tools and data sets based on the KB's digitised collection.

 5.482 lines of code

 56.330 MB files


 60 events

Datasets

more datasets 

Tools


<https://lab.kb.nl/>

more tools 



DBNL OCR Data set

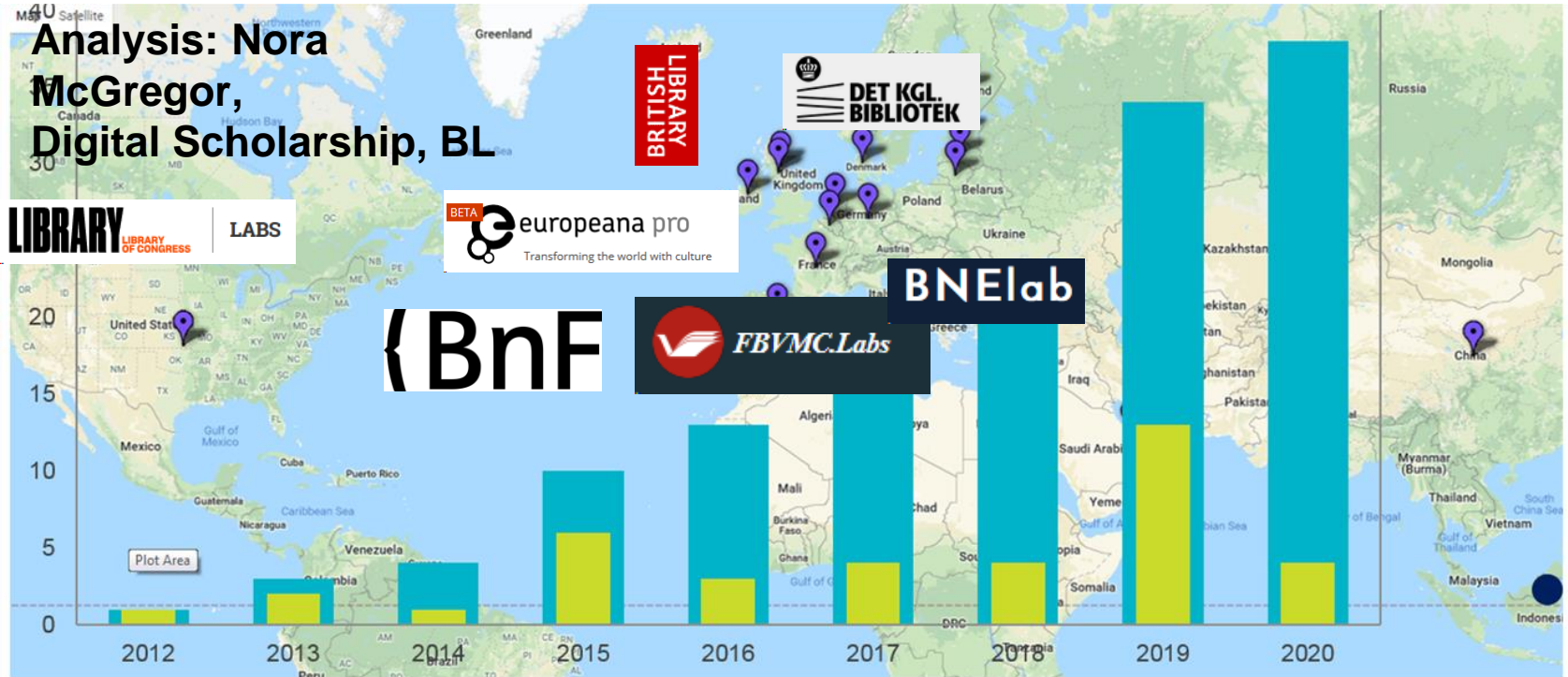


 RDA Entity Finder



 Brinkeys Tool

Growing 'Library' Labs International Community



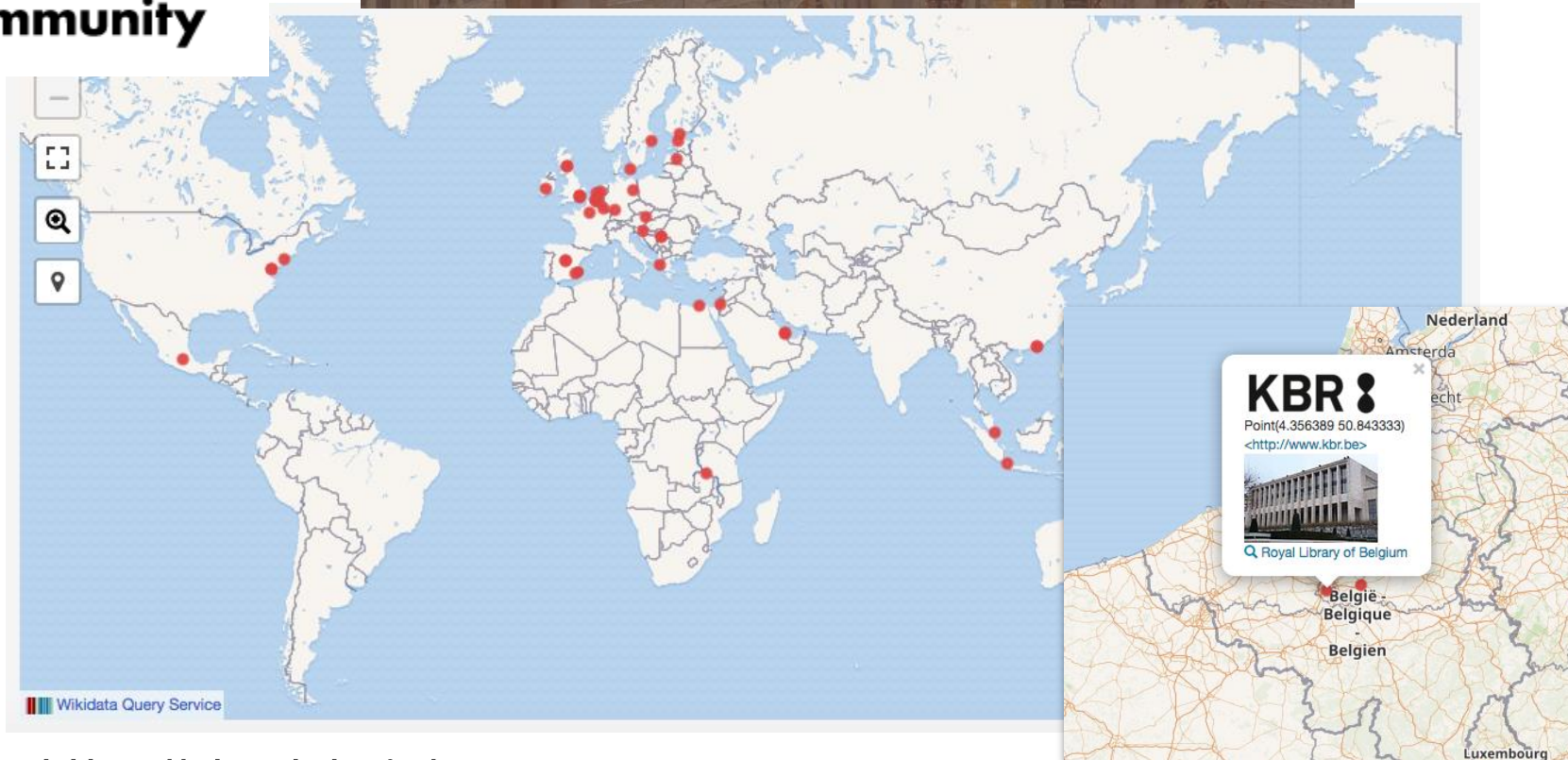
It's about people!



YouTube: <https://bit.ly/2YMxXYj>

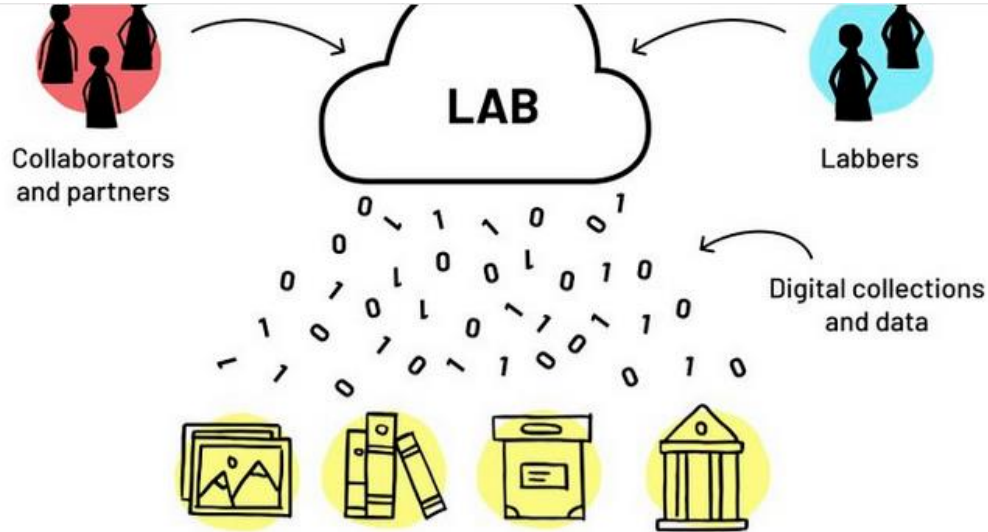
International GLAM Labs Community

A community of the world's innovative
cultural and research institutions



<https://glamlabs.io/>

International GLAM Labs Community



<https://glamlabs.io/books/>

1st KBR Labs event @ DATeCH 2019



1.data access: using APIs to access digital library collections

2.data preparation: preparing the data for analysis

3.data analysis: tools and methods for analysing your data

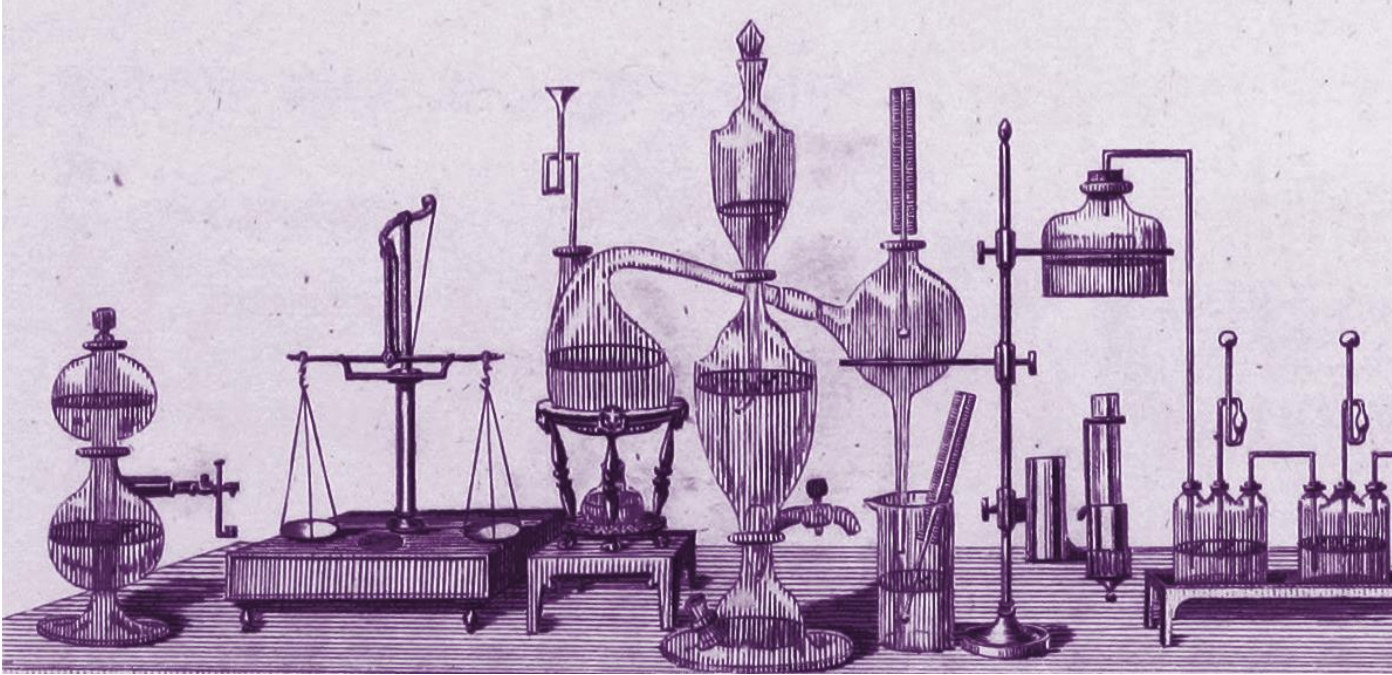
Experimenting in the Digital Research Lab: a hands-on introduction to using digital collections for digital humanities research



Royal Library of Belgium by Leandro Neumann Guiffo

Royal Library of Belgium, 8 May 2019, 09:00 – 18:30

<http://datech.digitisation.eu/programme/workshop/>



KBR 8



GHENT
UNIVERSITY

Towards the KBR Digital Research Lab



FED-**t**WIN

Towards the KBR Digital Research Lab

- FED-tWIN: programme of sustainable research cooperation between the Federal Scientific Institutions & Belgian universities
- “Towards a KBR Digital Research Lab and knowledge network for the study of Belgium’s digitised and born-digital historical, literary, publishing and cultural heritage of the 19th - 21st centuries.”
- To establish a long-term cooperation between the Ghent Centre for Digital Humanities (GhentCDH) and the Royal Library of Belgium (KBR) - 10 years initially

KBR 



**GHENT
UNIVERSITY**



FED-*t*WIN

Towards the KBR Digital Research Lab



The aim of the KBR Digital Research Lab is:

- facilitate data-level access to KBR's digitised and born-digital collections
- ensure that the digitised and born-digital collections are embedded into researcher's workflows in a user-friendly manner
- optimise the digitised collections for using digital humanities research methods, such text and data mining



Towards the KBR Digital Research Lab

KBR

Next step: recruit a postdoctoral researcher

Job Opportunity: Permanent Lecturer / Post Doctoral Researcher for the KBR Digital Research Lab

GhentCDH



KBR

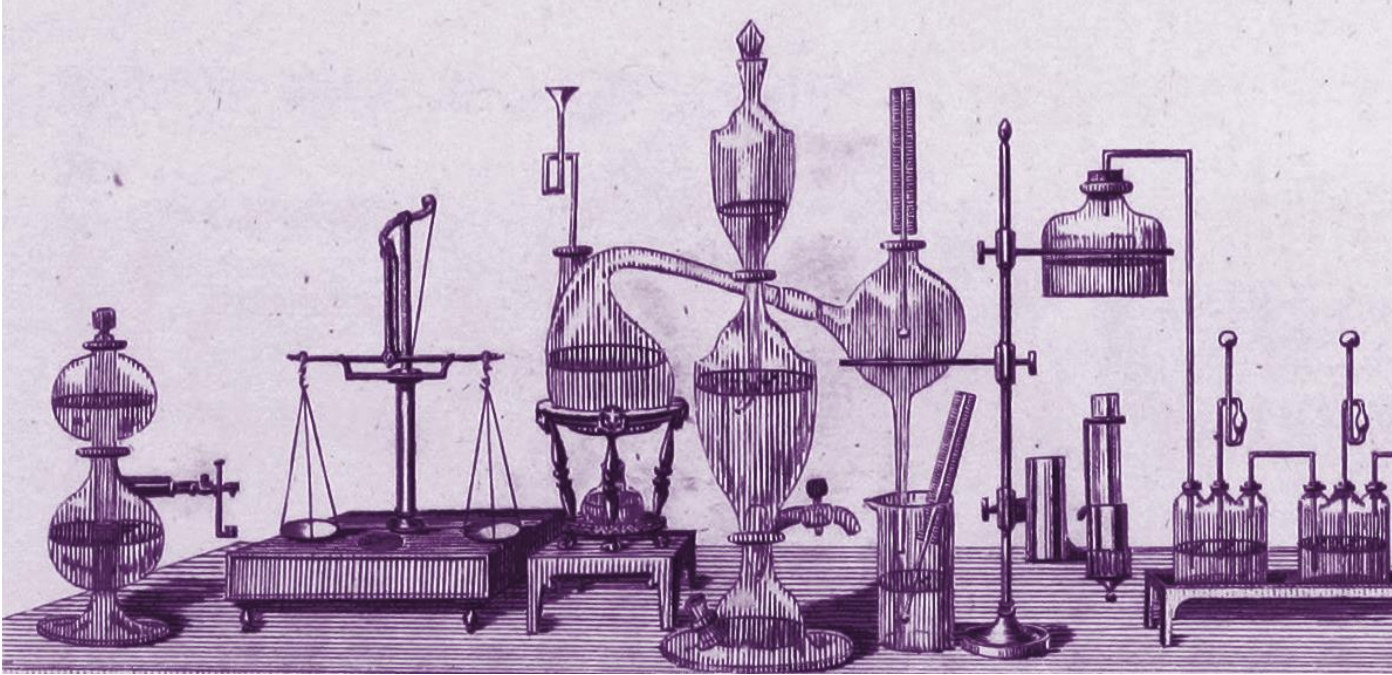


The Ghent Centre for Digital Humanities is delighted to welcome applications for a post doctoral researcher for the KBR Digital Research Lab. The aim of the KBR Digital Research Lab is to establish a long-term cooperation between the Ghent Centre for Digital Humanities and KBR, Royal Library of Belgium to facilitate data-level access to KBR's digitised and born-digital collections for digital humanities research.

[Read more](#)



FED-tWIN



KBR 8



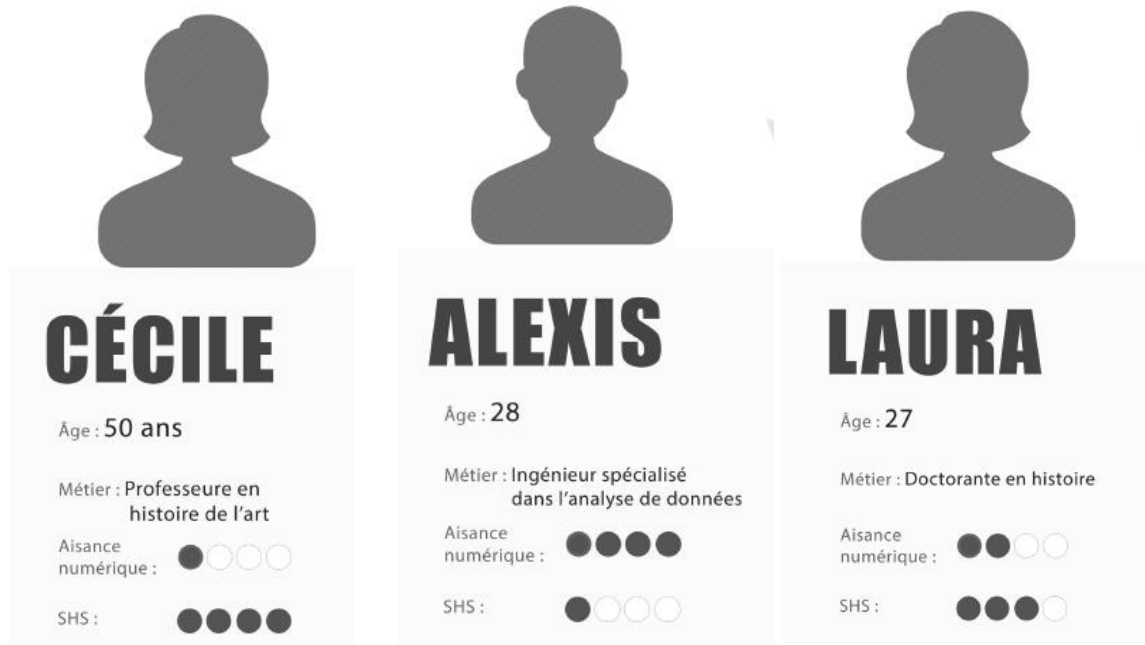
GHENT
UNIVERSITY

Understanding user needs



FED-**t**WIN

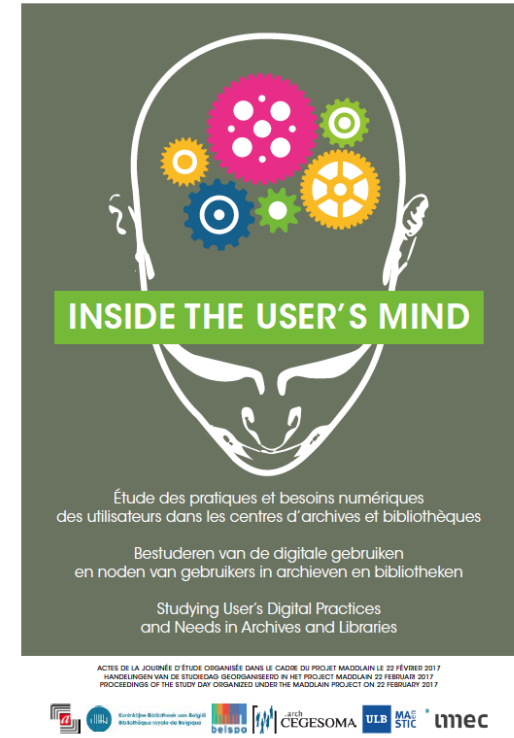
Researcher needs



Eleonora Moiraghi (2018) Le projet Corpus et ses publics potentiels. : Une étude prospective sur les besoins et les attentes des futurs usagers. [Rapport de recherche] Bibliothèque nationale de France. <https://hal-bnf.archives-ouvertes.fr/hal-01739730>

User needs: MADDLAIN

The aim of the MADDLAIN project was to analyse the behaviour and the needs of different audiences with regard to digital access to the collections offered by the documentation institutions of the Federal Science Policy (CegeSoma, KBR and State Archives of Belgium).



<https://www.kbr.be/en/projects/maddlain/>

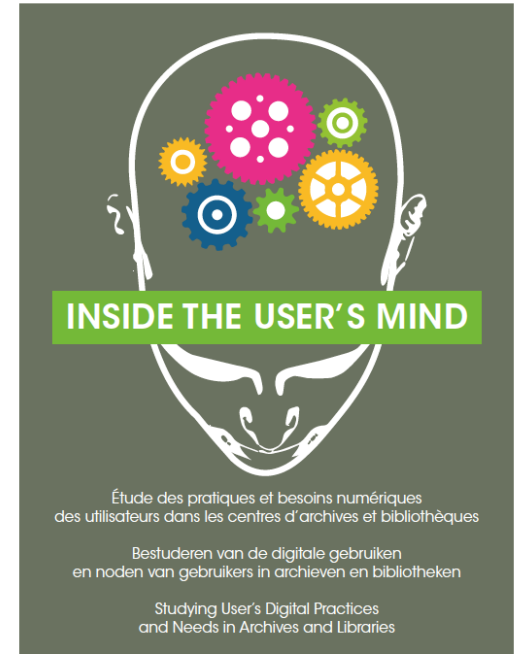
User needs: MADDLAIN

The project consisted of three components:

- A study concerning general access to digital information.
- A study regarding the use of e-learning by university students who are studying humanities subjects.
- A study focusing on the use of Virtual Research Environments by researchers in the humanities.

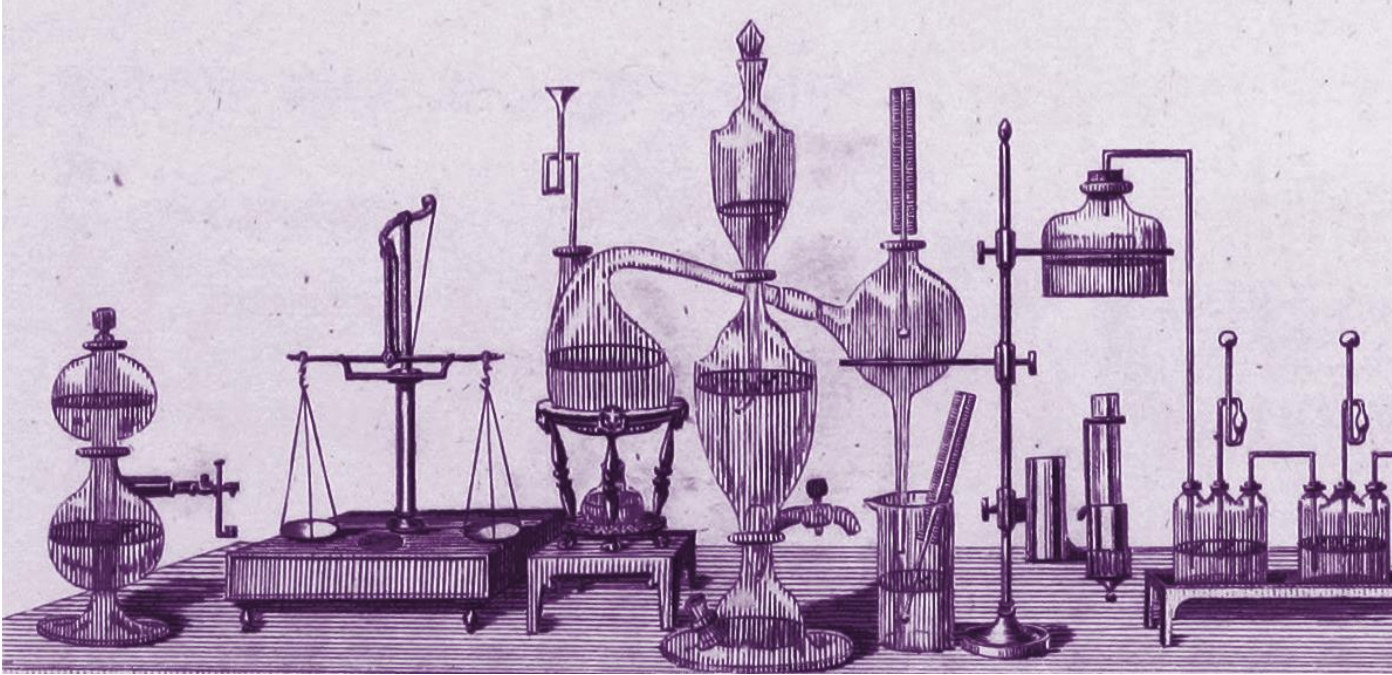
Final Report:

<https://opac.kbr.be/Library/doc/SYRACUSE/10160844/maddlain-identifying-needs-to-modernize-access-to-digital-dat-in-libraires-and-archives-final-report>



ACTES DE LA JOURNÉE D'ÉTUDE ORGANISÉE DANS LE CADRE DU PROJET MADDLAIN LE 22 FÉVRIER 2017
HANDELINGEN VAN DE STUDIEDAG GEORGANISEERD IN HET PROJECT MADDLAIN 22 FEBRUARI 2017
PROCEEDINGS OF THE STUDY DAY ORGANIZED UNDER THE MADDLAIN PROJECT ON 22 FEBRUARY 2017





Potential Research Pilots

KBR 



**GHENT
UNIVERSITY**

Belgian Science Policy Office



belspo

FED-*t*WIN

Overarching research focus

“Towards a KBR Digital Research Lab and knowledge network for the study of Belgium’s digitised and born-digital historical, literary, publishing and cultural heritage of the 19th - 21st centuries.”

KBR 



FED-*t*WIN

Digital Research Methods

KBR



GHENT CENTRE FOR DIGITAL HUMANITIES

Search

[HOME](#) [SERVICES](#) [PROJECTS](#) [ACTIVITIES](#) [ABOUT THE CENTRE](#)

The GhentCDH

The Ghent Centre for Digital Humanities (GhentCDH) engages in the field of 'Digital Humanities' at Ghent University, ranging from archaeology and geography to linguistics and cultural studies. It develops DH collaboration and supports research projects, teaching activities and infrastructure projects across the faculties.

Digital text analysis

The GhentCDH aims to improve digital text analysis at Ghent University by offering support and information to researchers. You can contact us for advice on TEI and digital editions, working with digital text analysis tools, and using computer-assisted qualitative data analysis.

[Read more about this service](#)

DH Support | Geospatial analysis | **Digital text analysis** | Collaborative databases



www.ghentcdh.ugent.be/services/digital-text-analysis

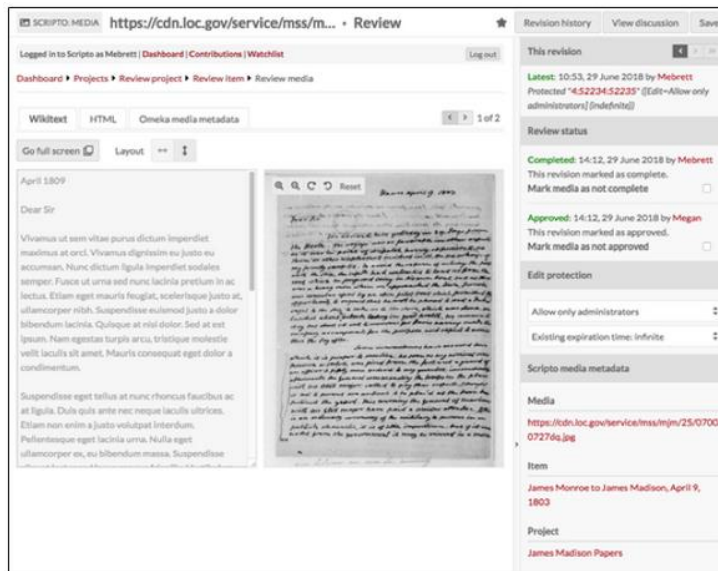


Towards a digital text analysis platform



KBR 8

- Annotations (e.g. crowd-sourced)
- Annotations (e.g. Named Entity Extraction)
- Correction of OCR
- Manual Transcription



www.ghentcdh.ugent.be/services/digital-text-analysis



Potential Research Pilots: Collective Action Belgium

- **Discipline:** social history
- **Period:** interbellum and WWII
- **Topic:** tracing the dynamics of contention, strikes, demonstrations and other forms of collective action in Belgium.
- **Methods:** Named Entity Recognition (NER) and Geographic Entity Recognition (GER)

KBR 



FED-*t*WIN

Potential Research Pilots: Opening up the Belgian Bibliography

- **Discipline:** library / data science
- **Period:** 1966-
- **Topic:** Opening up the bibliographic metadata of the Belgian bibliography for research use
- **Method:** Bibliographic transition of the Belgian Bibliography using the IFLA Library Reference Model (LRM) including FRBR (Functional Requirements for Bibliographic Records) and RDA (Resource Description and Access)

KBR 



**GHENT
UNIVERSITY**



FED-*t*WIN

Potential Research Pilots: The feuilleton in Belgium

- **Discipline:** literature
- **Period:** 1830-1930
- **Topic:** mapping the publication of literature in Belgian newspapers across the first century of the Belgian nation state.
- **Method:** image recognition to locate feuilleton sections in newspapers, extract the text for analysis with digital text analysis tools

KBR 

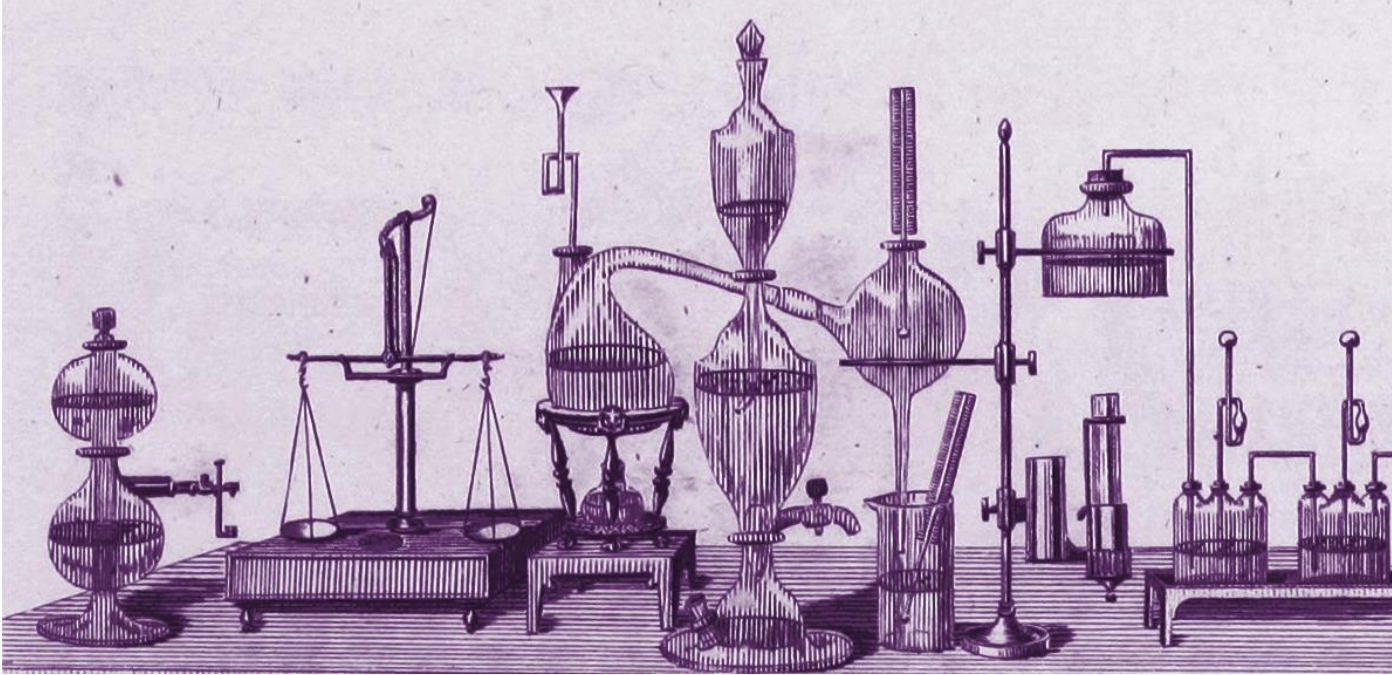


**GHENT
UNIVERSITY**

Belgian Science Policy Office



FED-*t*WIN



KBR 8



GHENT
UNIVERSITY

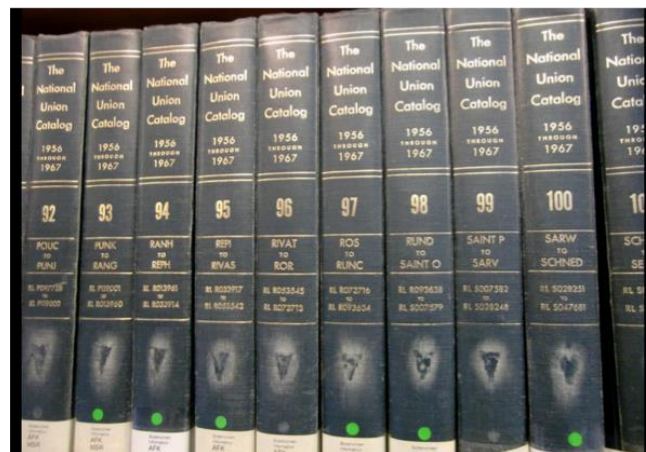
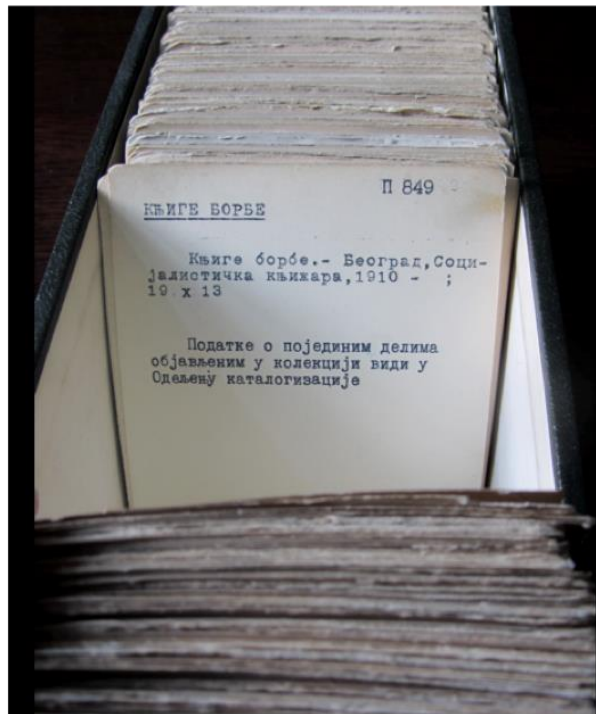
Data-level access to collections



FED-**t**WIN

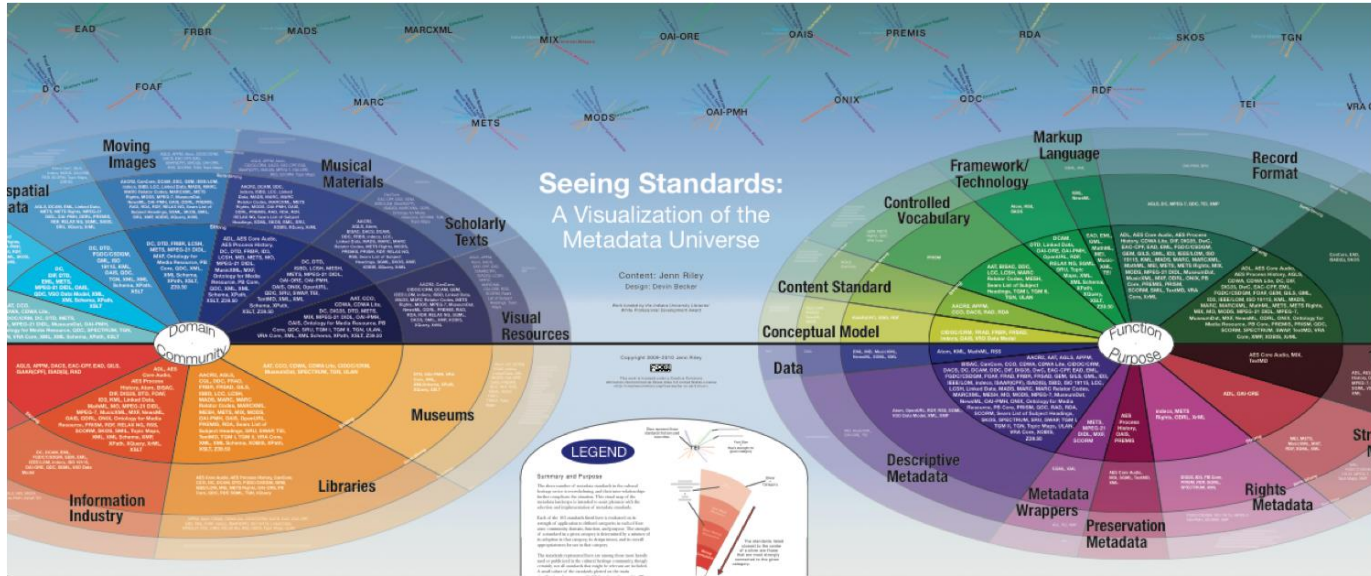
'Analogue' collections

KBR



FED-tWIN

Metadata



<http://jennriley.com/metadatamap/>



Digitised Newspapers



Historical Newspapers



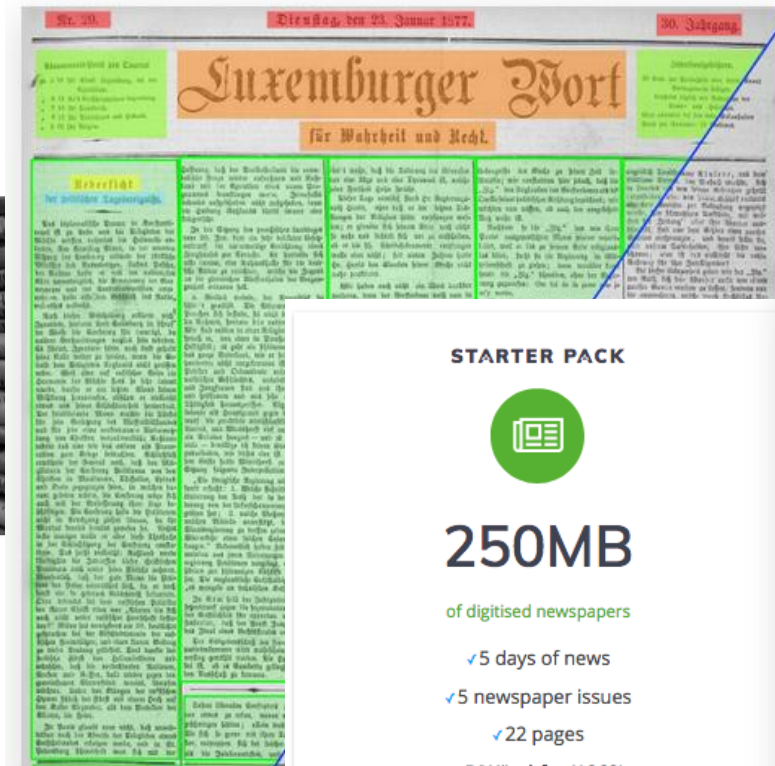
Open Standards



High Quality Data



Clear Copyright



STARTER PACK



250MB

of digitised newspapers

- ✓ 5 days of news
- ✓ 5 newspaper issues
- ✓ 22 pages
- ✓ D'Wäschfra (1868)
- ✓ Public Domain, CC0 (See copyright notice)
- ✓ Best for getting started & developing

 DOWNLOAD (ZIP)

<https://data.bn.l.lu/data/historical-newspapers/>

Digitised collections



EDUCATION RESEARCH FACILITIES WORKING AT UGENT LIVING IN GHENT ABOUT US INFORMATION FOR ∨

[Home](#) > [Facilities](#) > [University library](#) > [Google Books @ UGent](#)

Google Books @ UGent

In a nutshell

Google started digitizing the Booktower's out-of-copyright books by the end of 2007, bringing some of the greatest works of Dutch and French literature to millions of people globally. As one of Belgium's largest libraries, the Ghent University Library will greatly increase the number of Dutch and French-language books available through the Google Book Search Library Project. Through this partnership, hundreds and thousands of books will be added to the program, from well-loved classics to rare and special collections previously only available to those able to consult the library's stacks. Once the books have been digitized, it will be easy for anyone to discover, search and read them on the

- [Google Books](#)
- [UGent books in Google Books](#)
- [Google & UGent : press documentation \(May 23rd, 2007\)](#)
- [Google & UGent : press release \(May 23rd, 2007\)](#)
- [Google & UGent : press documentation \(August 28th, 2009\)](#)

<http://www.ugent.be/en/facilities/library/google-project.htm>

KBR



GHENT
UNIVERSITY

Belgian Science Policy Office



FED-*t*WIN

Hand-written documents

KBR

Transkribus

Register

Login

Transcribe. Collaborate. Share...

...and benefit from cutting edge research in Handwritten Text Recognition!

Download version 1.0

Download version 1.0 for Mac

Wiki » How-to guide (pdf) »

Scholars

Are you transcribing historical documents? Handwritten or printed, from the middle ages or from the 20th century? Would you like to do this in a highly standardized, flexible and reliable way? And do you appreciate to get support from automated tools such as Handwritten Text Recognition and Layout

The screenshot displays the Transkribus web interface. On the left, there is a sidebar with navigation options like 'Overview', 'Layout', 'Metadata', and 'Tools'. Below this is a 'Recently documents' section and a 'Collections' dropdown menu. A table lists documents with columns for ID, Title, Pages, and Uploader. The main area shows a document titled 'Sept 1824' with handwritten text. A circular stamp with the number '44' and '1824' is visible on the right side of the document. Below the document, there is a list of transcription suggestions with numbered markers (4, 5, 6, 7) and red arrows pointing to specific words in the text.

ID	Title	Pages	Uploader	Updated
4863	Bertham Workshop D...	18	berthamwork...	The Ap...
4857	Transkribus	2	berthamwork...	The Ap...

4 the individual states, being a imitation of the house of Peers
5 in Britain, is a separation of power without any utility and
6 perhaps only protracts legislation and is the cause of on-
7 necessary dissensions _ the senate in the federal Government



GHENT
UNIVERSITY

Belgian Science Policy Office



FED-tWIN

<https://transkribus.eu/Transkribus/>

Born-digital collections, e.g. web-archives

KBR

The screenshot shows the BnF website interface. At the top, there are navigation links for various languages (Español, Français, Deutsch, Italiano, Português, Русский, العربية, 中文, 日本語), utility links (Aide, sitemap, Write to BnF), and a user account link (Espace personnel). The BnF logo and name 'Bibliothèque nationale de France' are on the left. A search bar and 'Soutenez Richelieu' button are in the center. A red navigation bar contains categories: THE BnF, COLLECTIONS AND SERVICES, CULTURAL EVENTS, PROFESSIONALS, and RESOURCES. Below this is a large image of a Japanese woodblock print. The main content area shows a breadcrumb trail: Home > Collections and services > Collections by subjects > Book, press and media > Internet archives. The title 'Internet archives' is prominent. Below it, text states: 'Web archives collected by the BnF can be consulted in research reading rooms either at François-Mitterrand or Richelieu Library.' A list of bullet points follows: 'A legal framework established in 2006', 'Some representative samples', and 'Modes of consultation'. On the right, there is a 'CONTACT' section with an email address 'depot.legal.web@bnf.fr' and a 'FOR MORE INFO' section with a link to 'Digital legal deposit: four questions about Web Ar...'. A vertical sidebar on the right contains social media icons for RSS, Blog, Lettres, Facebook, Twitter, YouTube, and Instagram.

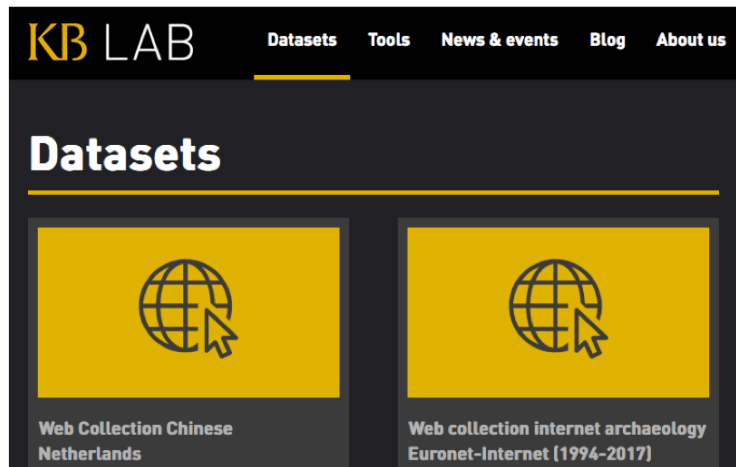


http://www.bnf.fr/en/collections_and_services/book_press_media/a.internet_archives.html

FED-**t**WIN

Datasets

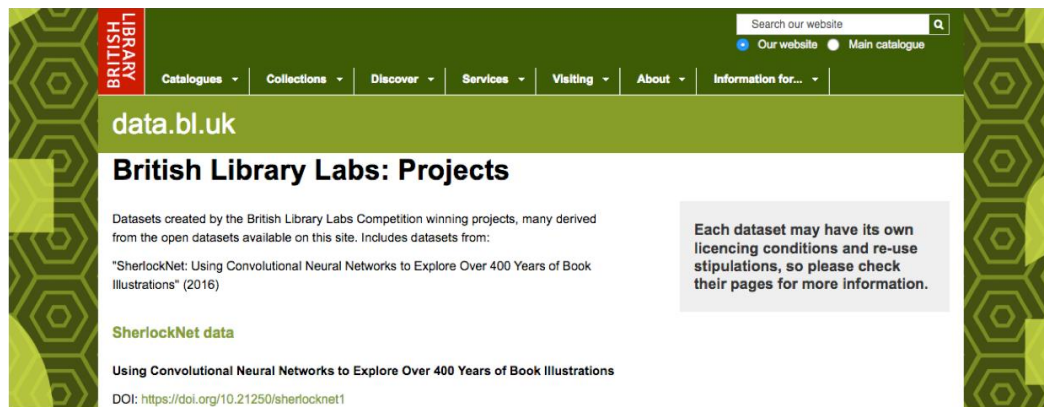
<http://lab.kb.nl/datasets>



KBR



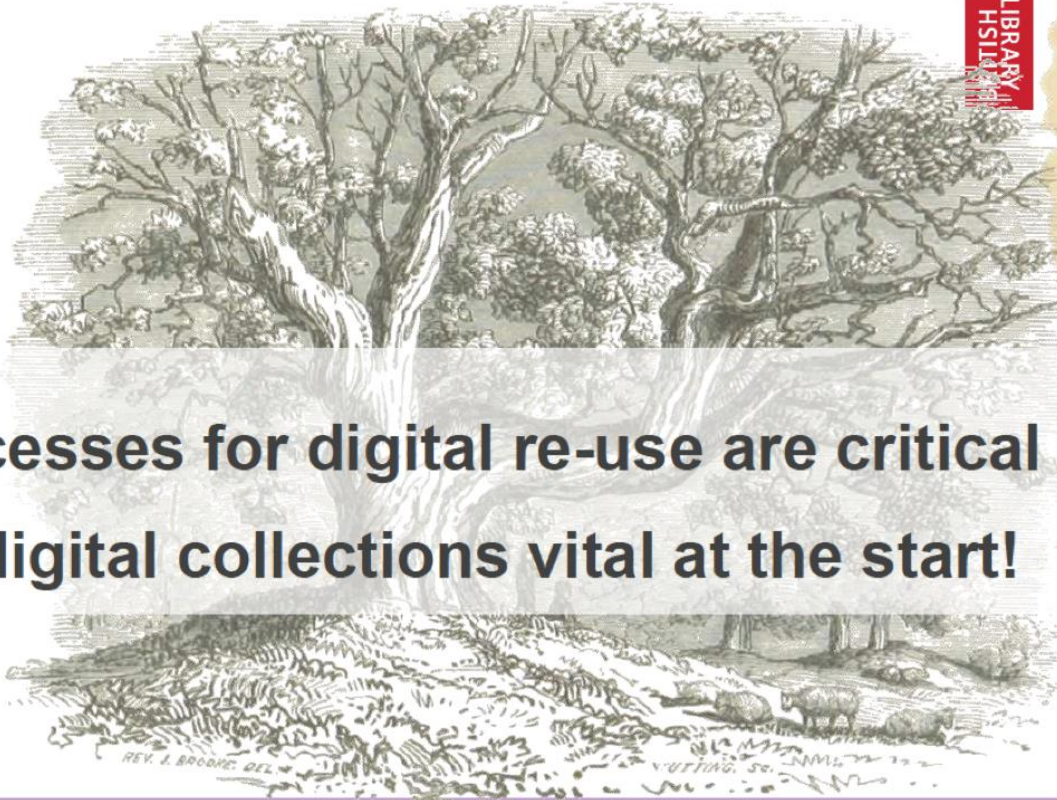
GHENT
UNIVERSITY



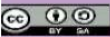
<https://data.bl.uk/research/>

FED-tWIN

'BL Labs' lessons 1



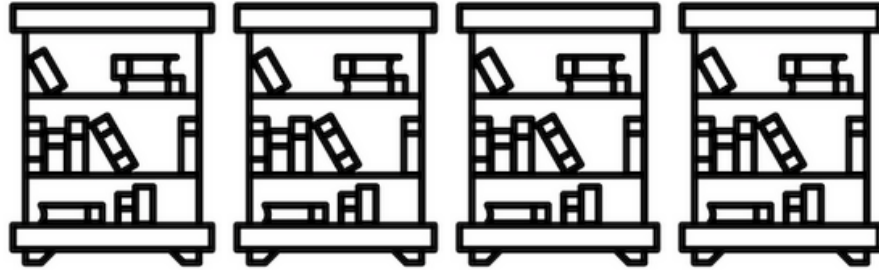
Policies and processes for digital re-use are critical
Having a list of digital collections vital at the start!



[@BL_Labs](#) [@mahendra_mahey](#) [@UCL_Qatar](#) [@britishlibrary](#) [@BL_DigiSchol](#) mahendra.mahey@bl.uk

With thanks to Mahendra Mahey, BL Labs

Collections as Data: “Always Already Computational” - Final Report



KBR 



Padilla, Thomas, Allen, Laurie, Frost, Hannah, Potvin, Sarah, Russey Roke, Elizabeth, & Varner, Stewart. (2019, May 22). Final Report --- Always Already Computational: Collections as Data (Version 1). Zenodo. <http://doi.org/10.5281/zenodo.3152935>



Always Already Computational: Collections as Data final report and project deliverables: <https://osf.io/mx6uk/wiki/home/>

FED-*t*WIN

www.kbr.be/en/collections/

Collections



Newspapers

The collection In terms of diversity and volumes, KBR's collection of newspapers is by far the [...]

Continue reading



Belgian Bibliography

The Belgian Bibliography makes inventories of catalogue records of all non-periodical publications (monographs, brochures, offline electronic [...])

Continue reading



Colloquium Saving the Web

30 October 2019

On 18 October 2019, the colloquium 'Saving the Web: the Promise of a Belgian Web Archive' [...]

Continue reading

Towards data.kbr.be



- **Project proposal:** “DATA-KBR-BE: Facilitating Data-level Access to KBR Collections for Open Science“
- **Aim:** to optimise KBR’s existing ICT infrastructure to stimulate sustainable FAIR (Findable, Accessible, Interoperable, Reusable) data-level access to KBR’s digitised collections for digital humanities research, i.e. to set-up **data.kbr.be**



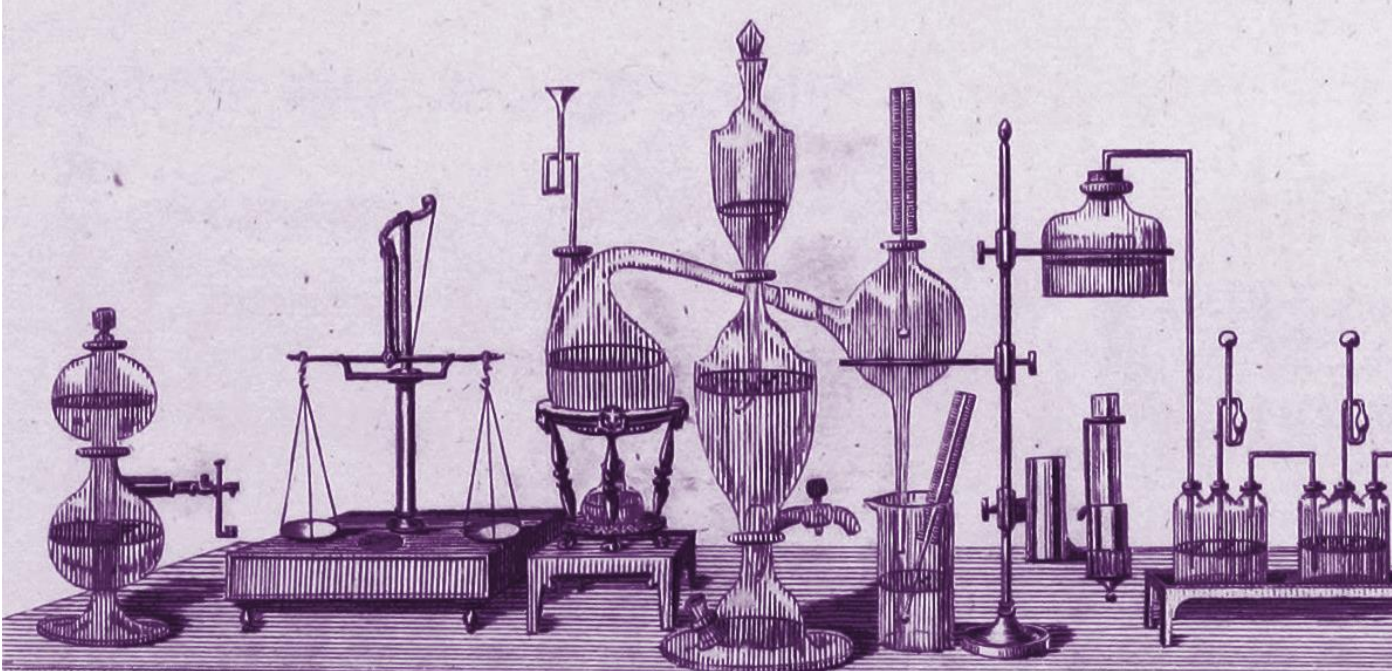
It's about people!



A graphic featuring a collection of stylized, colorful hands raised in the air. The hands are rendered in various colors including purple, orange, yellow, green, blue, pink, dark blue, teal, and light green. They are arranged in a dense, overlapping group, with some hands appearing in the foreground and others behind, creating a sense of depth and collective action. The background is plain white.

Come and join us

KBR 8



Merci! Bedankt! Thank you!

FED-*t*WIN