



Datensets Tools Kollektionen Artspace

<!-- ÖNB . Labs -->

Die ÖNB Labs sind die Plattform der Österreichischen Nationalbibliothek für digitale Bestände als Daten zum Forschen, Experimentieren und kreativen Verarbeiten.

Unter unseren [Datensets](#) finden Sie ausgewählte digitale Bestände der Österreichischen Nationalbibliothek, darunter Bilder, Volltexte oder Metadaten (Katalogdaten als Linked Open Data, Webarchivdaten). Eine Auswahl an [Tools](#) zur Verarbeitung unserer Daten finden Sie in unserer Tool-Section. Dort finden Sie Eintritt zu unserem GitLab für verschiedene Skripts und Projektdaten.

Sollten Sie Unterstützung für Ihr Projekt benötigen, wenden Sie sich bitte an uns. Die ÖNB Labs verstehen sich als dynamisches Serviceangebot der Österreichischen Nationalbibliothek, das laufend nach den Anforderungen der Community erweitert und verändert wird.

Die folgende Galerie präsentiert Beispielprojekte, welche die vielfältigen Einsatzmöglichkeiten unseres Angebotes zeigen sollen. Wir laden Sie ein, [mitzumachen](#) und Ihre Projekte mit der Labs Community zu teilen.



Österreichische
Nationalbibliothek

5+ Years of ONB Labs: Looking Back, What's New, and Looking Ahead

Christoph Steindl &
Johannes Knüchel

ONB Labs History

- Part of the ONB strategy for Digital Humanities
- Conceptual design and development since 2017
- Launch 2018
- Three main components
 - Topics
 - Datasets
 - Tools

Why are the ONB Labs so important?

ANNO

AKON

ALMA

Situation Now

ONB [' Labs ']



Data formats



GitLab

Infrastructure



Interactive
technologies



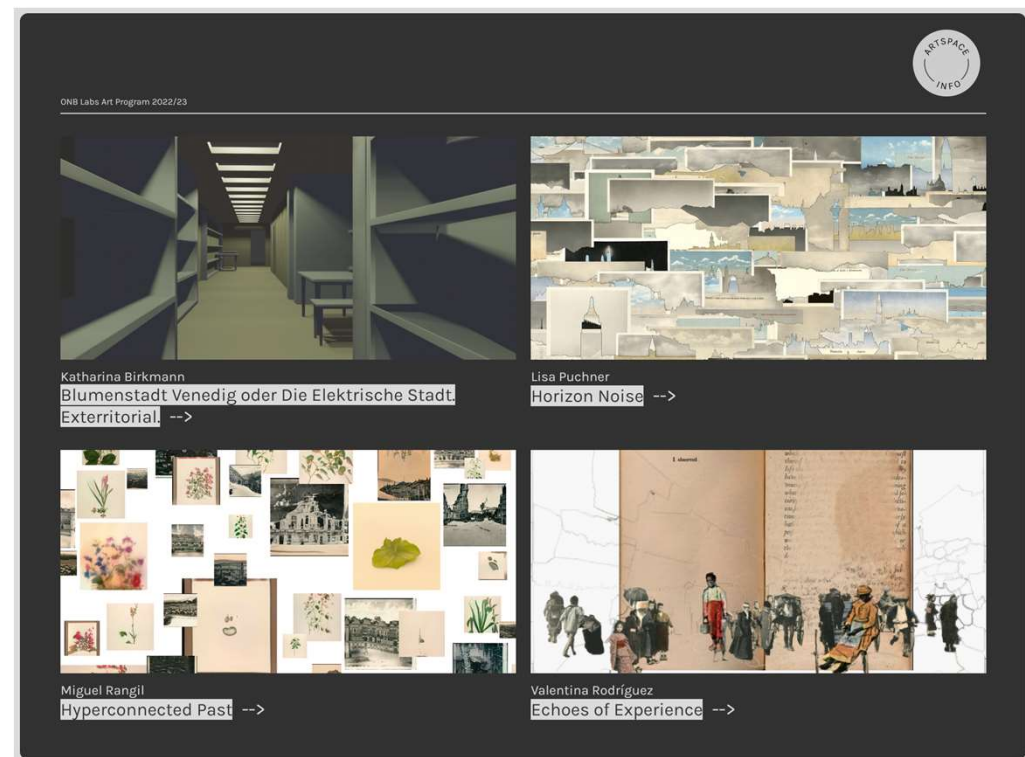
Events and Activities

- 3 ONB Labs symposiums
- 3 art programs
- Methodological course at University of Vienna
- Participations
 - Booksprint "Open a GLAM Lab"
 - Pydays
 - Internships
 - Activities in GLAM sector
 - 2 Hackathons



ONB Labs and Art

- 3 art programs
 - 2× online and 1× physical
- ONB Labs Artspace



ONB Labs and Newspapers

- Provides access to newspaper data older than 140y
- IIIF technology
- Dataset is often requested by scientific use cases
- OCR data available
- Part of research projects
 - Impresso
 - NewsEye
 - ...

New Data – New Stories

- ONB Labs are part of new digitisation initiatives
 - Strengthen sustainability aspect
 - Improve dissemination of results
- New technologies in future?
- Need to contextualise data

What's New

- New Data Sets
- Improved Data Sets
- Improved Documentation of Certain Data Sets
- Glossary
- Bilingual ONB Labs

Looking Ahead

- Results of the Labs Symposium 2024
- Results of Community Efforts
- Expanding our Offer of Jupyter Notebooks
(CLARIAH-AT funded Project 2025)

Wrap-Up Day 1





Österreichische
Nationalbibliothek

Thanks for coming,
let's have some food
and wine!

New Data Set

ESPERANTISTS MEET KING REPRESENTATIVE

Registration for Congress to
Open Monday Exceeds
Previous Records.

AUTHOR OF TONGUE
IS EXPECTED TODAY

Capt. Josefo Perogordo, of Spain,
Comes as Representative
of Alfonso.

Final preliminaries for the formal
opening of the Esperantists' congress,
which will take place Monday evening,
were made today by the local commit-
tee in charge at Arlington Hotel head-
quarters.

Arrangements for the business ses-
sion have been completed, an enter-
taining social program has been an-
nounced, and everything is in readiness
for Monday's opening. Virtually all
arrangements had been made by this after-
noon, when the delegates went in
large numbers to Union Station to
meet Dr. L. L. Zamenhof, of Warsaw,
and, the author of the new inter-
national language.

It was announced at headquarters to-
day that when registrations and the
exchange of tickets for the congress
were counted up, it will be found that
previous records for attendance at
Esperantist conventions will have been
broken.

Envoy of Alfonso.

Among those who registered today one
of the most notable figures is Capt.



DR. L. L. ZAMENHOF,
Of Warsaw, Poland Author of
Universal Tongue.

Esperanto Newspaper Excerpts

Historical newspaper
clippings (from the period
1898-1915) of the Hachette
Collection, containing
information about the
Esperanto language and
community

22247 JPG files

17204 articles

22 languages

New Data Set



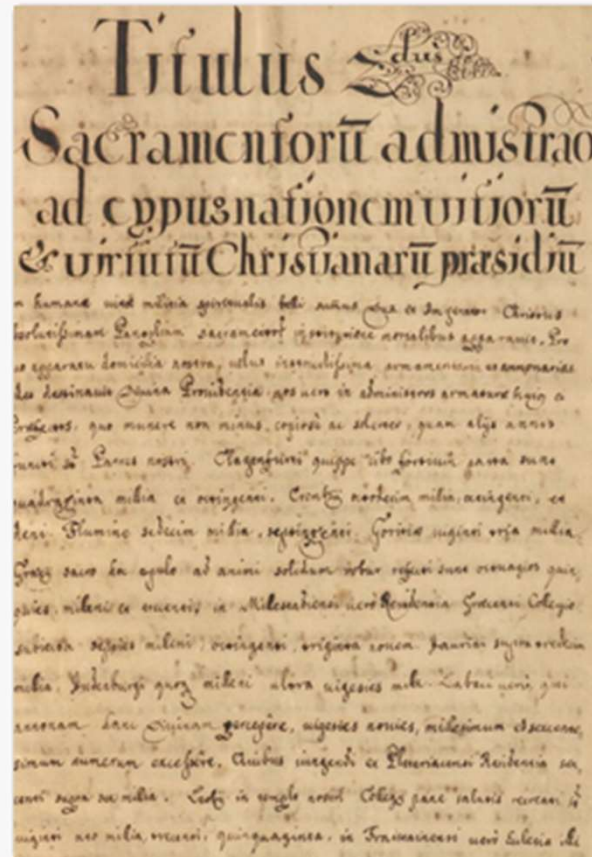
Johann Caspar Lavater Collection

Scans of hand drawings
and prints including
metadata from the
graphics collection of
Johann Caspar Lavater
(1741-1801)

21903 IIIF Manifests

21903 JPG files

New Data Set



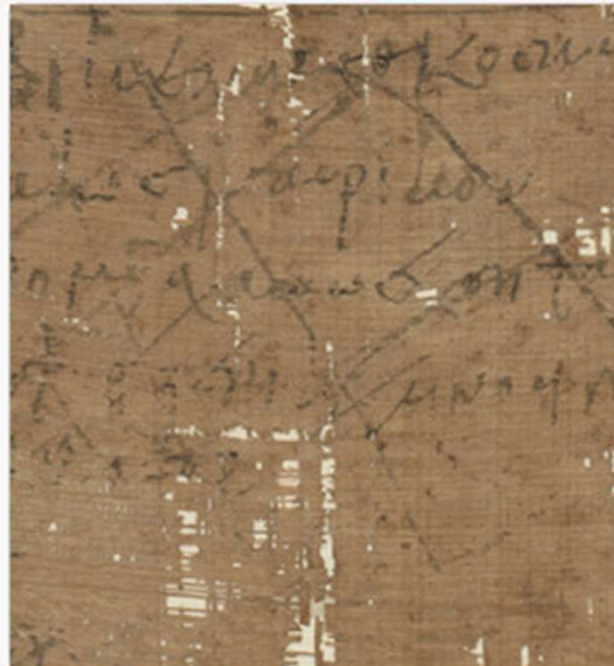
Jesuit Chronicles

Images, bibliographical and codicological metadata for the Corpus Jesuiticum, containing information on the Jesuit order in Central Europe

485 manuscripts

138505 images

Improved Data Set



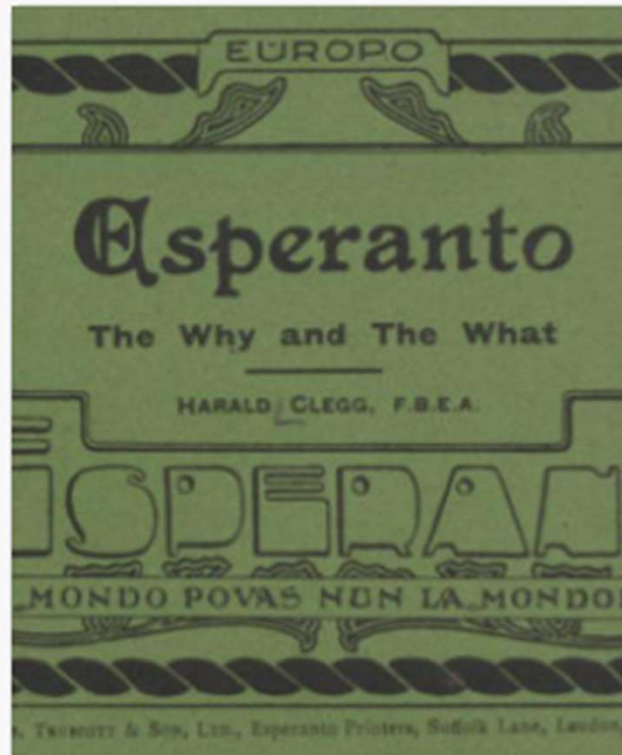
Papyri

Digital objects of the
Papyrus Museum of the
Austrian National Library

15108 documents

10 languages

Improved Data Set



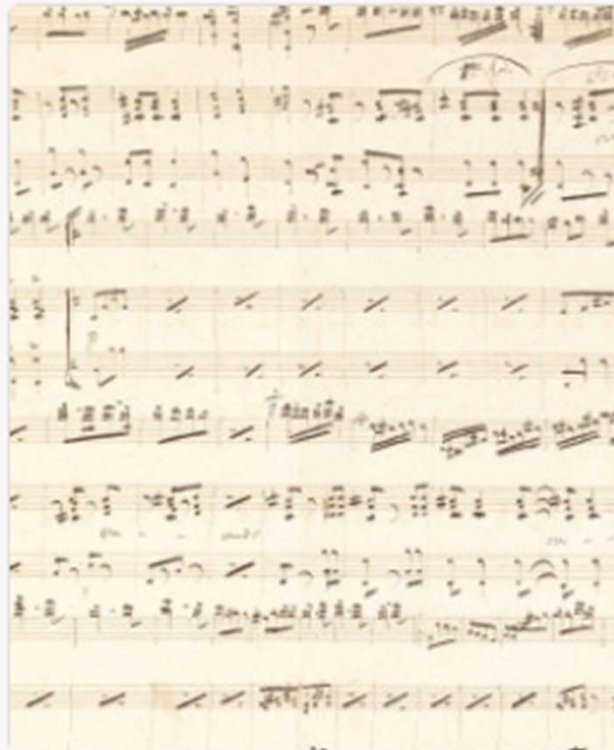
Planned Languages

Digitized objects from the
Department of Planned
Languages of the Austrian
National Library

1957 documents

40 languages

Improved Data Set



Musical Manuscripts

Digitized Musical
Manuscripts of the
Department of Music of
the Austrian National
Library

1251 composers

4205 manuscripts

Improved Data Set to come



This data set is currently unavailable due to the documents being redigitized.

Pamphlets 1848

Pamphlets from the 1848 revolution in the Habsburg Monarchy

3206 JPG files

9 languages

Improved Documentation of Data Sets

Maintenance

This data set can be updated irregularly.

Use Cases

Areas of Application Related to the Data Set

Possible Uses

The images in the data set could be used to train a CV[●]/ML[●]-model that is meant to be able to automatically recognize landscapes and places. Since the data set contains metadata to image documents, it could be used to gain an exemplary understanding of the distribution of places over time, and to therefore gain an insight into certain characteristics that might be present in the entire collection.

Examples

IIIF and ML

Project on IIIF and Machine
Learning Using the Historical
Postcards Data Set

ML

Coloring Historical Postcards
Using Machine Learning

GAN

Generating Synthetic Postcards
Using a Generative Adversarial
Network

Glossary

The images in the data set could be used to train a HTR¹-model. They could also be used to apply a CV¹-model that can recognize different styles of writing within the written sources, potentially leading to insight about the different writers of the manuscripts.

Glossary

This table provides a comprehensive overview of the technical terms and abbreviations. Here, you will find both the German as well as the English terms and definitions to ensure clear and consistent use of the terminology.

Term	Definition
ABO	<i>Austrian Books Online</i> . Project (and associated collection) of the → ONB, in which it digitizes its historical book holdings in a <i>Public Private Partnership</i> with <i>Google</i> . See https://www.onb.ac.at/en/digital-offers/austrian-books-online
AC number	<i>Austrian Central Catalog</i> number. → OBV-wide identifier of the bibliographic metadata record after formal indexing; corresponds to an abstract description of all copies, so one AC number can correspond to several copies or barcodes

Bilingual Labs

(: \$ ÖNB // Labs :)

de en

ONB Labs – A Timeline



KJ0

vielleicht als Option interessant

KNÜCHEL Johannes; 2024-11-18T13:39:40.850



