



JAGIELLONIAN UNIVERSITY  
IN KRAKOW

---

# **Digitalization of library handwriting resources and opportunities for Artificial Intelligence. The case of Autographa Sammlung in the Jagiellonian Library**

---

8/12/2022

Roman Sosnowski, Piotr Tylus, Jadwiga Kita-Huber, Remigiusz Sapa,  
Grzegorz J. Nalepa, Krzysztof Kutt  
Jagiellonian University



Photo: Paul Korecky CC-BY-SA 2.0

## Historical overview

- The collection originates from the former Prussian State Library (Preussische Staatsbibliothek)
- 1661 Churfürstliche Bibliothek zu Cölln an der Spree (Elector's library) – since 1701 Königliche Bibliothek (Royal Library) – since 1918 Preußische Staatsbibliothek (Prussian State Library)
- Now it is: Staatsbibliothek zu Berlin - Preußischer Kulturbesitz

## The way to Kraków

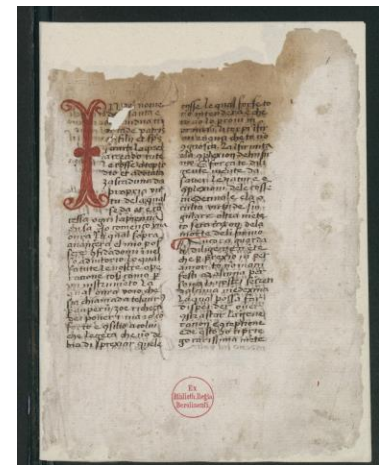
1941 – valuable collections transported out of Berlin, part of them in Lower Silesia (Fürstenstein today named Książ)

1943 – moved to Grüssau (today named Krzeszów)

1945 – collection secured and then moved to Kraków (Delegate Dr Stanisław Sierotwiński)

As a result, now, some subcollections are split between Kraków and Berlin (e.g. Ms. Gallica), some subcollections are entirely in Berlin or in Kraków.

Often the split collections were studied in their entirety: e.g. the medieval French manuscripts: D. Stutzmann, P. Tylus, *Les Manuscrits médiévaux français et occitans de la Preussische Staatsbibliothek et de la Staatsbibliothek zu Berlin–Preussischer Kulturbesitz*, Wiesbaden 2007.





BIBLIOTEKA

JACEŁŁŃSKA

WIEŚCIE



## Main parts of the collection

490 boxes. The presence of the collection in the Jagiellonian Library remained largely unknown until the mid-1980s.

1. Musical collections and graphical collections
2. Bound manuscripts: Americana, Gallica (218 units), Hispanica (111 units), Germanica, Graeca, Italica (150 units) and Lituanica, Latina, Latina theologica, also Arabic, Syrian, Indian and Aramaic manuscripts.
3. "Collection Varnhagen" and the "Collection Autographa" with a specific (alphabetical organization) – very large collections referring and embracing an important part of the cultural heritage of Europe especially of the 18<sup>th</sup> and 19<sup>th</sup> centuries.
4. Personal archives: Jakob Michael Reinhold Lenz, Wilhelm and Alexander von Humboldt, Hoffmann von Fallersleben, Gustav Freytag and George Schweinfuhrer.
5. East Asian Manuscripts and books (more than 2000 units),  
including the Beijing Collection, Pander collection, Hirth, Möllendorf and Müller, Libri Sinesi and Libri coreanici.

## What has been done so far? NOT DIGITALLY

### **Monika Jaglarz (Jagiellonian University), Katarzyna Jaśtał (Jagiellonian University)**

German manuscripts and archives. The general structure of the collection

### **Zdzisław Pietrzyk (Jagiellonian University)**

Polonica in Albums forming part of the former Prussian State Library in Berlin

### **Paweł Zarychta (Jagiellonian University)**

Rosa Maria und David Assings in Varnhagen collection

### **Jadwiga Kita-Huber (Jagiellonian University)**

Female Writers in the Varnhagen Collection – Letters, Works, Relations

### **Ottmar Ette (Universität zu Potsdam), Jutta Weben (Staatsbibliothek zu Berlin - Preußischer Kulturbesitz), Dominik Erdmann (Universität zu Berlin), Monika**

### **Jaglarz (Jagiellonian University), Ulrike Leitner (other researchers)**

Alexander von Humboldt

### **Piotr Tylus, Roman Sosnowski, Anna Rzepka, Jadwiga Miszańska (other researchers) (Jagiellonian University)**

Romance manuscripts in the Berlin collection at the Jagiellonian

### **Jacek Partyka (Jagiellonian University)**

Polonica and Polish books from the collection of the former Prussian State Library in Berlin now in the Jagiellonian Library in Krakow

### **Paul Videsott (Libera Università di Bolzano / Freie Universität Bozen) Giorgio Cadorini (Slezská univerzita v Opavě)**

Ladin and Friulan manuscripts (Böhmer papers)

### **Jan Stradomski (Jagiellonian University)**

Slavic Manuscripts

### **Katarzyna Kołakowska (Katolicki Uniwersytet Lubelski)**

Greek Manuscripts

### **Julia Madajczak (Uniwersytet Warszawski)**

Nahuatl manuscripts

### **Marek Mejor (Uniwersytet Warszawski), Thupten Kunga Chashab (Uniwersytet Warszawski)**

Pander collection and Tibetica

### **Carlo Pulsoni (Università degli Studi di Perugia)**

Romance manuscripts and Aldus Manutius' printed books

### **Nikolaus Gatter (Varnhagen Gesellschaft)**

Ludmilla Assing

### **Gesa Weinert (Universität Bonn)**

J. M. R. Lenz in the former Prussian State Library

### **Jana Kittelmann (Martin-Luther-Universität Halle-Wittenberg)**

Manuscripts of Johann Georg Sulzers in Varnhagen collection and Autographs collection

### **Jochen Strobel (Universität Marburg)**

August Wilhelm Schlegel's correspondence

### **Ingo Breuer (Universität zu Köln)**

Letters of Heinrich von Kleist

### **Heinz Härtl (Arnim-Arbeitsstelle der Stiftung Weimarer Klassik)**

Letters from Varnhagen collection

### **Lesław Łesyk (Katolicki Uniwersytet Lubelski)**

The literary tradition of Kourtesios in the former Prussian State Library in Berlin now in the Jagiellonian Library

### **Izabela Surynt (Uniwersytet Wrocławski)**

Gustav Freytag's letters



## Autographa Sammlung

196 boxes with the folders in the alphabetical order (from A to Hirzel and from Kromayer to Z)

Separately boxes with the legacies/archives of Alexander von Humboldt, Paul Heyse, Heinrich von Kleist. Sometimes thematic boxes

Autographa similar to Varnhagen Collection – so the similar scheme of description (set of metadata) will be adopted



Photo: Paweł Zarychta

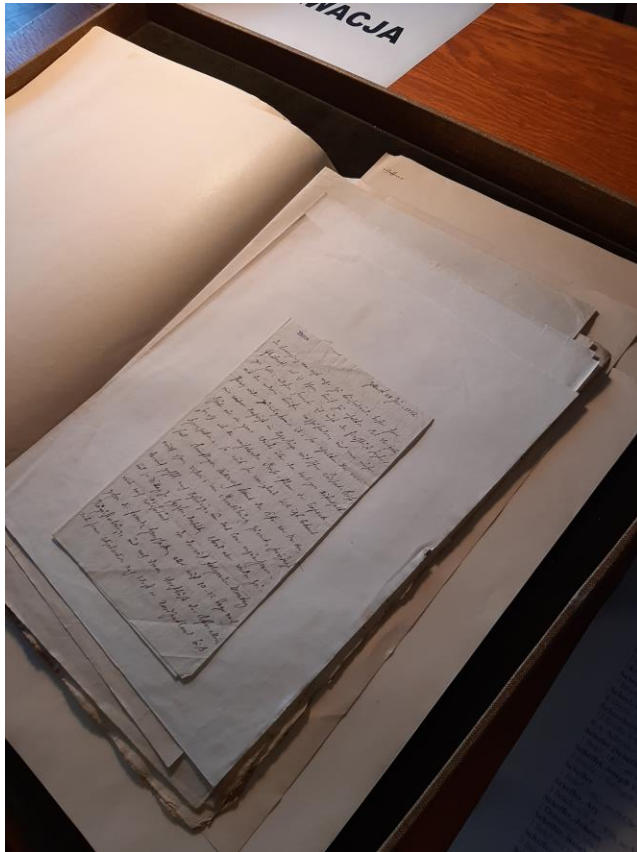




Photo: Paweł Zarychta



JAGIELLONIAN UNIVERSITY  
IN KRAKOW





# DESCRIPTION SCHEME FOR THE AUTOGRAPH COLLECTION

Signature of the whole unit (SA, name, surname)

1. Paintings and drawings
2. Outgoing correspondence
3. Creative works (literary and similar)
4. Personal material
5. Historical material, diplomas
6. Prints and press cuttings
7. Incoming correspondence
8. Extraneous material (notes on the person, letters from various to various)
9. Library materials (wrappers, bookmarks)

[File Word with the detailed description](#) (Monika Jaglarz)

# Varnhagen Collection - 11 letters of K.A. Varnhagen to H. v. Chézy

Before foliotation

SV 47 (general signature)

After foliotation i.e. after numbering leaves (and after dividing into subunits with proper references).

- 8.37 Varnhagen Karl August, Berlin, 18.11.1835 fol. 731-732
- 8.38 Varnhagen Karl August, Berlin, 11.11.1847 fol. 733-734
- 8.39 Varnhagen Karl August, Berlin, 31.05.1852 fol. 735-736
- 8.40 Varnhagen Karl August, Berlin, 18.01.1854 fol. 737-738
- 8.41 Varnhagen Karl August, Berlin, 07.10.1854 fol. 739-740
- 8.42 Varnhagen Karl August, Berlin, 30.10.1854 fol. 741-742
- 8.43 Varnhagen Karl August, Berlin, 25.02.1855 fol. 743-744
- 8.44 Varnhagen Karl August, Berlin, 01.06.1855 fol. 745-746
- 8.45 Varnhagen Karl August, Berlin, 13.12.1855 fol. 747-748
- 8.46 Varnhagen Karl August, Berlin, 21.01.1856 fol. 749-750
- 8.47 Varnhagen Karl August, Berlin, 25.01.1856 fol. 751-752

## Focus on some subunits (SA outgoing correspondance; SA Buff, Lotte)

Seq.: 2.1

Folios: ff. 7-8

Language: French

Language: German

Contents: Letter of Lotte Buff to Rosa Varnhagen, on f. 8v annotation of Robert Buff

Sender: Frankenstein, Charlotte de domo Buff

Sender: Buff, Robert

Addressee: Assing, Rosa de domo Varnhagen

Place: Frankfurt

Date: 1807

Provenance: acc. ms. 1897.14

Edition: *Rosa Assings Briefwechsel*, Berlin 2000 p. 287

Notes: on f. 7v dried flower (glued)

Observations:

Seq.: 2.2

Folios: ff. 9-10

Language: German

Contents: Transcription of the Lotte Buff's letter to her mother

Nadawca: Frankenstein, Charlotte z domu Buff, 1o v. Kestner

Adresat: Buff, Luise de domo Weiss

Place: Weimar

Date: 05.02.1850

Provenance: Hegel

Edition:

Notes:

Observations:

Major  
infrastructure  
components

# Jagiellonian Digital Library

- [Link to JDL](#)
- A digital library run for 12 years by the Jagiellonian Library (the main library of the Jagiellonian University in Kraków) in order to both protect and make available to the public antique and valuable collections of the Jagiellonian Library. It currently contains and allows access to over 850,000 digital objects. It is based on the dLibra software commonly used by Polish digital libraries

Major  
infrastructure  
components

# Jagiellonian Digital Library

JDL provides:

- access to scans of objects converted to PDF
- application of OCR
- a separate collection with subcollections
- proper links and relationships between files
- secure file archiving

Major  
infrastructure  
components

# Jagiellonian University Library Catalogue

- [Link to JU Library Catalogue](#)
- The Jagiellonian University Library Catalogue makes it possible to find resources from the most of JU libraries in a single unified search, including: modern printed books and serials, old prints and manuscripts, graphic and cartographic objects, music prints, sound recordings, electronic documents, motion pictures, microfilms and others
- In addition, it belongs to a common, unified space that allows simultaneous search in several catalogs of selected large Polish libraries, including the National Library



Major  
infrastructure  
components

# Jagiellonian University Library Catalogue

## JU Library Catalogue:

- provides simplified metadata sets for all the objects scanned in a project (the metadata sets will be downloaded automatically from the project database)
- makes it possible to find the objects from the space of the library catalogue and transferring the user to the appropriate files in JDL
- promotes access to these resources and their visibility in the growing common network of large Polish libraries using ALMA

196 boxes, folders in the boxes, each folder = SIGNATURE (SA Name, Surname)

1. Librarians prepare the folders for every researcher (folders distribution by competence of each researcher – language or other specific competence)

2. The researcher opens the folder, analyses its contents, reads. Then reorganizes the documents in the folder, then makes foliation. Then fills the records with data (ARCARIUM). E.g. for a letter it will be (as in an example shown before):

Seq.: 2.1 / Folios: ff. 7-8 / Language: French / Language:

German / Contents: Letter of Lotte Buff to Rosa Varnhagen, on f. 8v annotation of Robert Buff / Sender: Frankenstein, Charlotte de domo Buff, / Sender: Buff,

Robert / Addressee: Assing, Rosa de domo Varnhagen / Place:

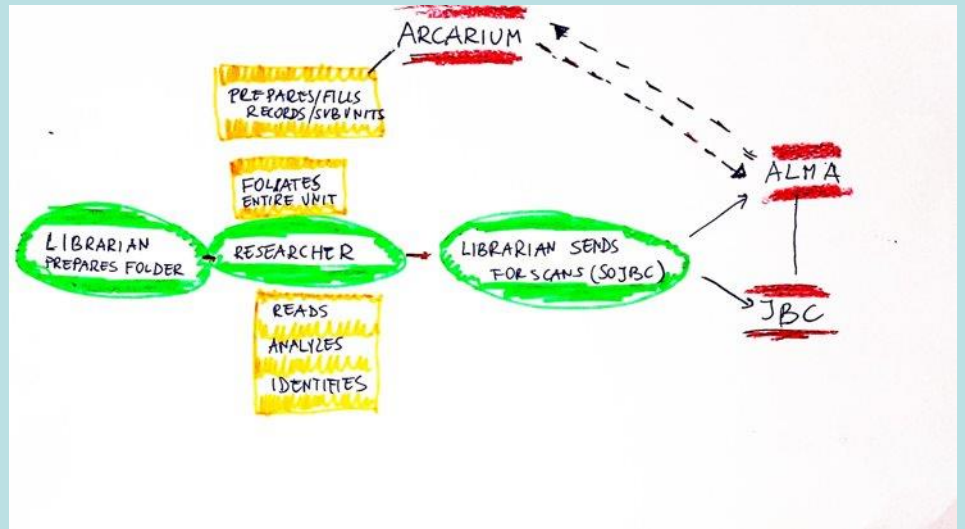
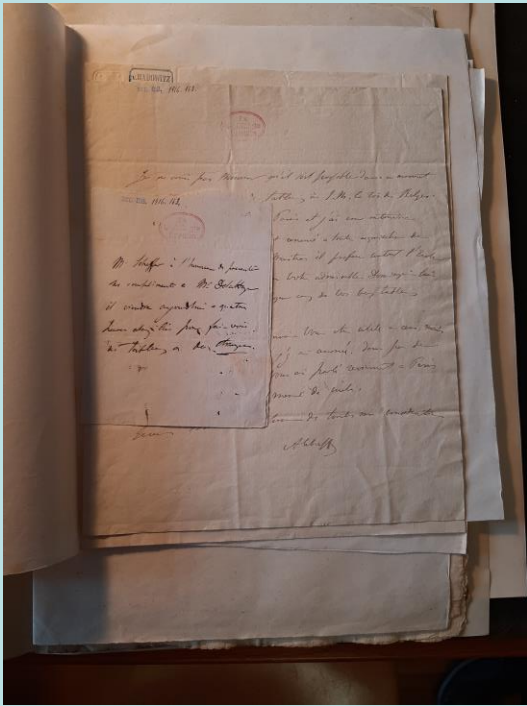
Frankfurt / Data: 1807 / Provenance: acc. ms. 1897.14 / Edition: *Rosa Assings Briefwechsel*, Berlin 2000 p. 287 / Notes: on f. 7v dried flower (glued) /

Observations:

ATTENTION! In a unit (folder/signature) there are more subunits (letters/records)

3. Librarians prepare the unit for scans (SOJBC), using data from previously prepared records

4. Digitization procedure (ALMA, JBC)



We have a rich collection of manuscripts containing many precious documents

We have effective and established digitalization procedures in JDL

**So, what are the challenges?**



# Record in ALMA and JBC

LIBRARY SEARCH BROWSE SEARCH BIBLIOTEKA JAGIELLOŃSKA DATABASES OF UJ DATABASES OF CMUJ JOURNAL SEARCH ...


991007835689705067 X Library Catalog ADVANCE


that contain my query words anywhere in the record

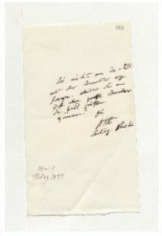
Sign in to get complete results and to request items Sign in X DISMISS

BACK TO RESULTS LIST Showing list of versions

0 selected 2 Results

1  MANUSCRIPT **Briefe an Eduard Buschmann Alexander von Humboldt.**  
Berol. Ms. Nachlass von Alexander von Humboldt, 13/2.; Humboldt, Alexander von (1769-1859);  
Buschmann, Johann Carl Eduard (1805-1880)  
1843-1848  
[Available Online](#)

2  MANUSCRIPT **Briefe an Eduard Buschmann Alexander von Humboldt.**  
Berol. Ms. Nachlass von Alexander von Humboldt, 13/1.; Humboldt, Alexander von (1769-1859);  
Buschmann, Johann Carl Eduard (1805-1880)  
1833-1842  
[Available Online](#)



Show content

Download

DESCRIPTION INFORMATION

Title:  
Brief von Alexander von Humboldt an Johann Carl Eduard Buschmann

Creator:  
Humboldt, Alexander von (1769-1859) ; Buschmann, Johann Carl Eduard (1805-1880)

Date:  
1847.03.14/1847.03.15

Resource Type:  
rekopis

Contributor:  
Buschmann, Johann Carl Eduard (1805-1880)

Description:  
DE-611-HS-302862 ; Nachlass Alexander von Humboldt Bd. 13/2 Bl. 392-393

Source:  
Biblioteka Jagiellońska, Berol. Ms. Nachlass von A. v. Humboldt, 13/2

Digital copy identifier:  
NBJGORP015864

Language:  
ger

Rights:  
Domena publiczna (public domain)

Location of original document:  
Biblioteka Jagiellońska

Relation:  
Katalog Kalliope

Computer Catalogue:  
click here to follow the link

Format:  
application/pdf

Resource Identifier:  
oai:jbc.uj.edu.pl:357076

Sponsorship:  
Biblioteka Państwowa w Berlinie

**We need to store more detailed data than stored in JBC and ALMA**

**We need different views**

**We need to sort and to connect the subunits (individual letters or individual documents).**

**We need a UNIT/SUBUNIT structure explicitated**

What is important from the textual point of view is to have a view on the whole library unit as well as on the subunits (individual letters/documents).

## Different systems/databases that deal with somehow similar material

Kalliope <https://kalliope.staatsbibliothek-berlin.de/>

The Wittgenstein Archives at the University of Bergen

(WAB) <http://wab.uib.no/sfb/> + <http://www.wittgensteinsource.org/>

Editionenportal Thüringen <https://www.editionenportal.de/>

Goethe- und Schiller-Archiv Weimar <https://ores.klassik-stiftung.de/ords/f?p=401:1:.....>

Jagiellonian Digital Library (JDL-JBC) <https://jbc.bj.uj.edu.pl/>

Fibula <http://info.filg.uj.edu.pl/fibula/en/>

Europeana <https://www.europeana.eu/en>

LetterSampo (Historical Letters on the semantic web): <https://lettersampo.demo.seco.cs.aalto.fi/en/>

Creation of the new database ARCARIUM (storing metadata)

Integration with the workflow of digitalization in Jagiellonian Library

Collection of data (ARCARIUM)

Longer perspective (but it is the task of AI specialists):

Tool that summarizes the contents of the manuscript automatically (for this we need AI powered handwriting recognition and AI powered multilingual summarizing tool).



## Structure of the project

FP consists of three platforms:

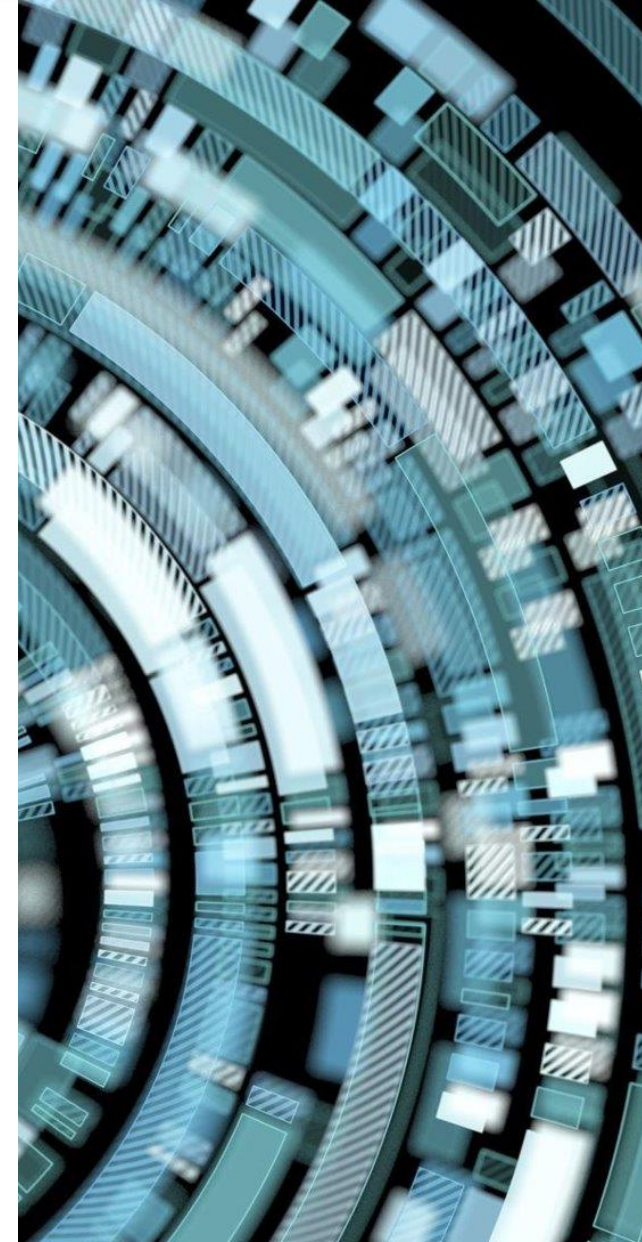
1. R2R platform related to the study of "Autographs" and the integration and stimulation of research on "Berlinka" [prof. Piotr Tylus];
2. R2R and R2B platform for implementation of Digital IT and Artificial Intelligence solutions [prof. Grzegorz J. Nalepa];
3. Infrastructure Platform for the preservation and digitization of resources [prof. Remigiusz Sapa].

The main subject is the Autographs Collection and each platform contributes to a better, scientific study of this collection.

## Structure of the project: Platform Two (G.J.Nalepa)

Platform Two includes three main threads in the areas of:

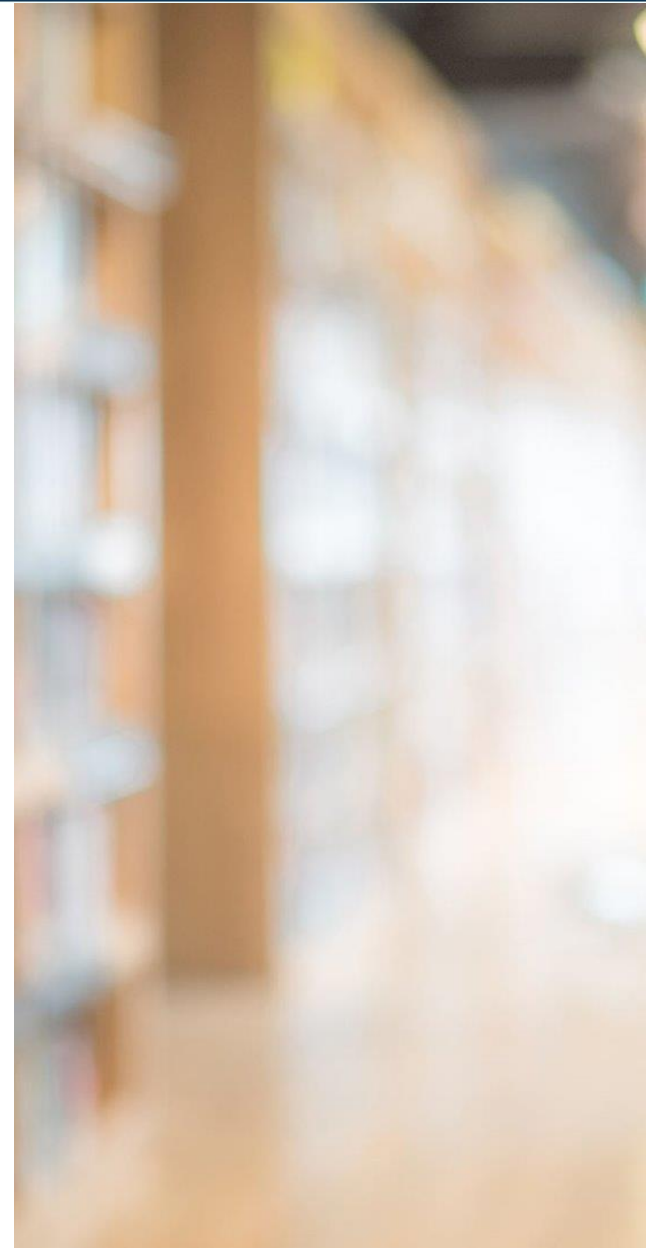
- Knowledge engineering and semantic technologies,
- AI-based solutions for literature studies,
- Cooperation with business and social environment interested in exploitation of results.



## Structure of the project: Platform Two

Objectives in the area of *knowledge engineering*:

1. Designing a knowledge engineering process supporting the digitalization procedures
2. Metadata acquisition and modelling
3. Development of ontological model supporting Arcarium
4. Proof of concept implementation of Arcarium
5. Integration with JU facilities
6. Integration with existing digital library systems for literature studies





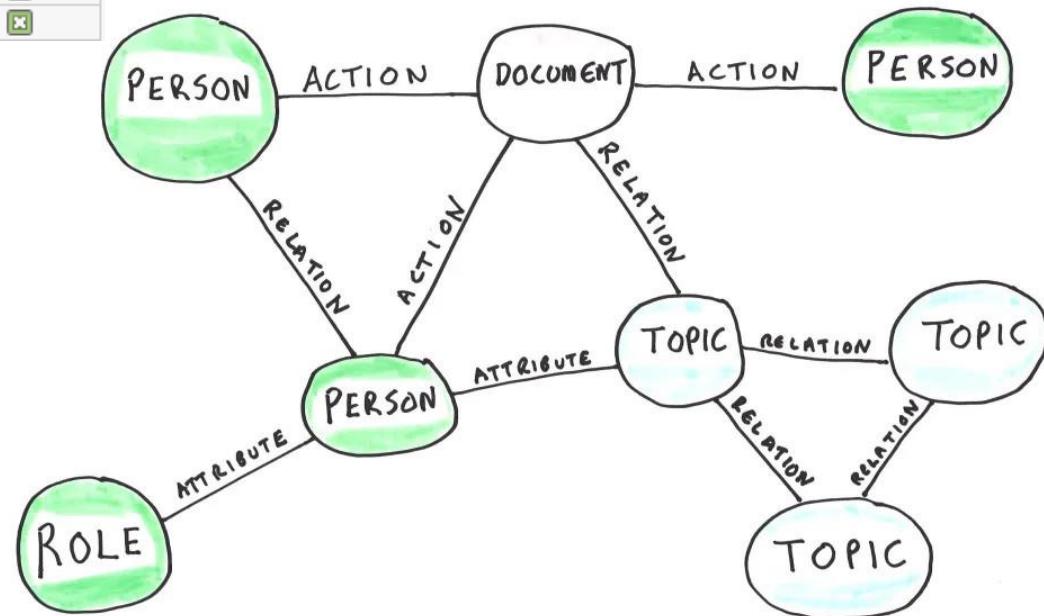
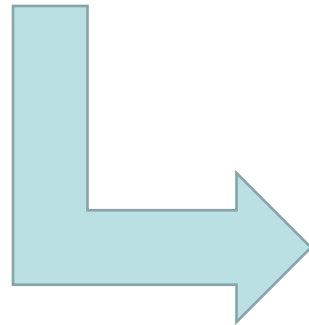
## Rich metadata as a foundation

ID	Title	ISBN	Author	Publishing...	isAvailable
1	1984	2343454895456	George Orwell	04/12/1979	<input checked="" type="checkbox"/>
2	Anna Karenina	1234548485843	Leo Tolstoy	07/11/1998	<input checked="" type="checkbox"/>
4	The Adventures of Huckleberry Finn	3450345345443	Mark Twain	08/11/1999	<input checked="" type="checkbox"/>
5	Ulysses	9944933003232	James Joyce	06/05/2010	<input checked="" type="checkbox"/>
8	War and Peace	0944344903312	Leo Tolstoy	08/11/2001	<input type="checkbox"/>
11	The Brothers Karamazov	9003940397271	Dostoievsky	04/07/2012	<input checked="" type="checkbox"/>
12	On the Road	0459450444310	Jack Kerouac	30/12/2005	<input type="checkbox"/>
15	The Metamorphosis	2003948930545	Franz Kafka	09/03/1976	<input checked="" type="checkbox"/>
16	The Illiad	9449039333923	Homer	05/07/1998	<input checked="" type="checkbox"/>
17	The Odyssey	8409404850139	Homer	06/08/1999	<input checked="" type="checkbox"/>



## Rich metadata as a foundation

ID	Title	ISBN	Author	Publishing...	isAvailable
1	1984	2343454895456	George Orwell	04/12/1979	<input checked="" type="checkbox"/>
2	Anna Karenina	1234548485843	Leo Tolstoy	07/11/1998	<input checked="" type="checkbox"/>
4	The Adventures of	3450345345443	Mark Twain	08/11/1999	<input checked="" type="checkbox"/>
5	Ulysses	9944933003232	James Joyce	06/05/2010	<input checked="" type="checkbox"/>
8	War and Peace	0944344903312	Leo Tolstoy	08/11/2001	<input type="checkbox"/>
11	The Brothers Karan	9003940397271	Doso	04/07/2012	<input checked="" type="checkbox"/>
12	On the Road	0459450444310	Jack Kerouac	30/12/2005	<input type="checkbox"/>
15	The Metamorphosi	2003948930545	Franz Kafka	09/03/1976	<input checked="" type="checkbox"/>
16	The Illiad	9449039333923	Homer	05/07/1998	<input checked="" type="checkbox"/>
17	The Odyssey	8409404850139	Homer	06/08/1999	<input checked="" type="checkbox"/>



## Rich metadata as a foundation makes a major *qualiative change*





JAGIELLONIAN UNIVERSITY  
IN KRAKOW

## Challenges...

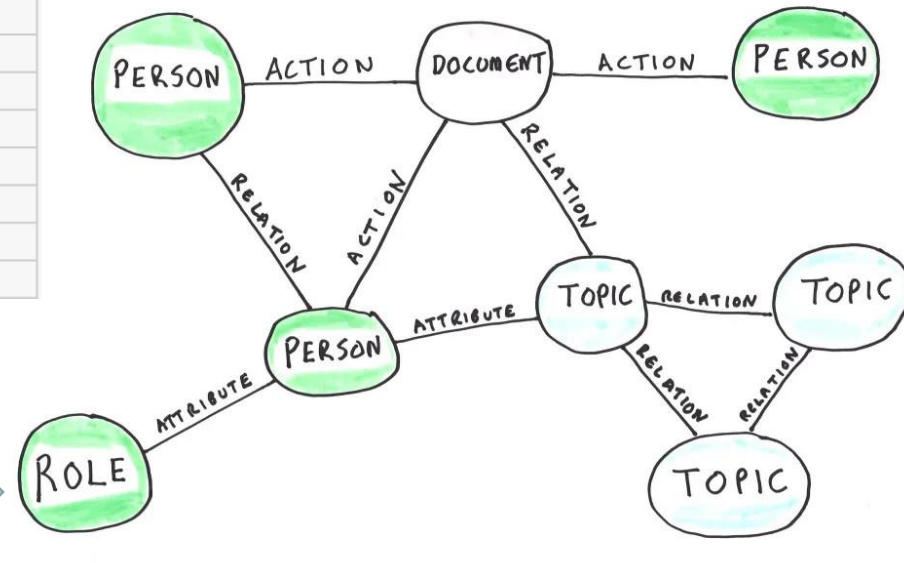
## Challenges

1) How to create a graph from tabular data?

*Are these two entries the same entity?*

*Do both entries follow the same date format?*

ID	Title	ISBN	Author	Publishing...	isAvailable
1	1984	2343454895456	George Orwell	04/12/1979	<input checked="" type="checkbox"/>
2	Anna Karenina	1234548485843	Leo Tolstoy	07/11/1998	<input checked="" type="checkbox"/>
4	The Adventures of I	3450345345443	Mark Twain	08/11/1999	<input checked="" type="checkbox"/>
5	Ulysses	9944933003232	James Joyce	06/05/2010	<input checked="" type="checkbox"/>
8	War and Peace	0944344903312	Leo Tolstoy	08/11/2001	<input type="checkbox"/>
11	The Brothers Karan	9003940397271	Doso	04/07/2012	<input checked="" type="checkbox"/>
12	On the Road	0459450444310	Jack Kerouac	30/12/2005	<input type="checkbox"/>
15	The Metamorphosi	2003948930545	Franz Kafka	09/03/1976	<input checked="" type="checkbox"/>
16	The Illiad	9449039333973	Homer	05/07/1998	<input checked="" type="checkbox"/>
17	The Odyssey	8409404850139	Homer	06/08/1999	<input checked="" type="checkbox"/>

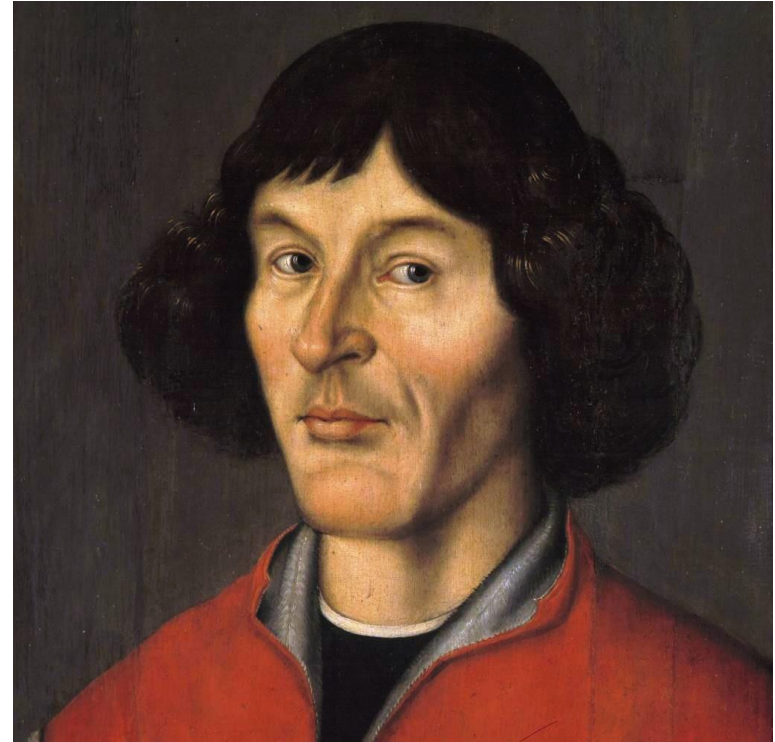




## Challenges

- 1) How to create a graph from tabular data?
- 2) How can AI methods help to identify and correct errors (which are unavoidable)?

*Are Nikolaus Kopernikus and Nikolaus Copernicus the same person?*

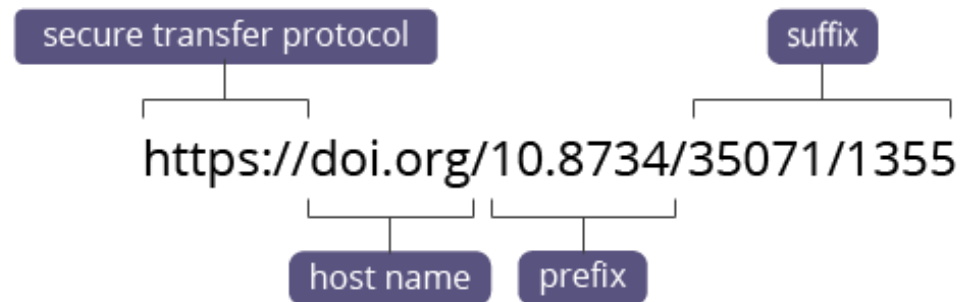


## Challenges

- 1) How to create a graph from tabular data?
- 2) How can AI methods help to identify and correct errors (which are unavoidable)?
- 3) How to prepare a **usable** hierarchy of identifiers to clearly distinguish units, documents and individual parts/pages?

*They must be unique, but also convenient for researchers to use (e.g., in research papers).*

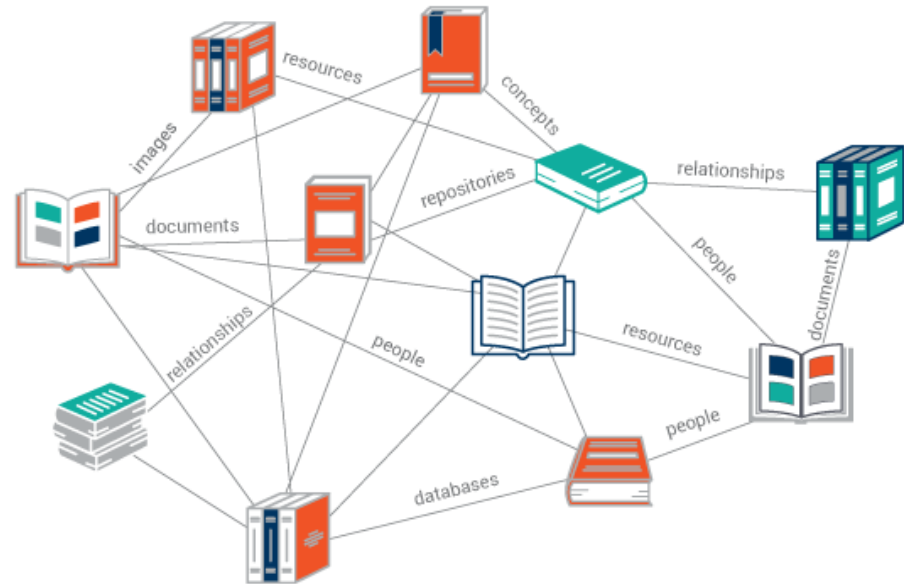
*URLs vs DOIs?*



## Challenges

- 1) How to create a graph from tabular data?
- 2) How can AI methods help to identify and correct errors (which are unavoidable)?
- 3) How to prepare a **usable** hierarchy of identifiers to clearly distinguish units, documents and individual parts/pages?
- 4) How can (semi-)automated modules enrich metadata, e.g., search and link other edits / comments / elaborations for each document?

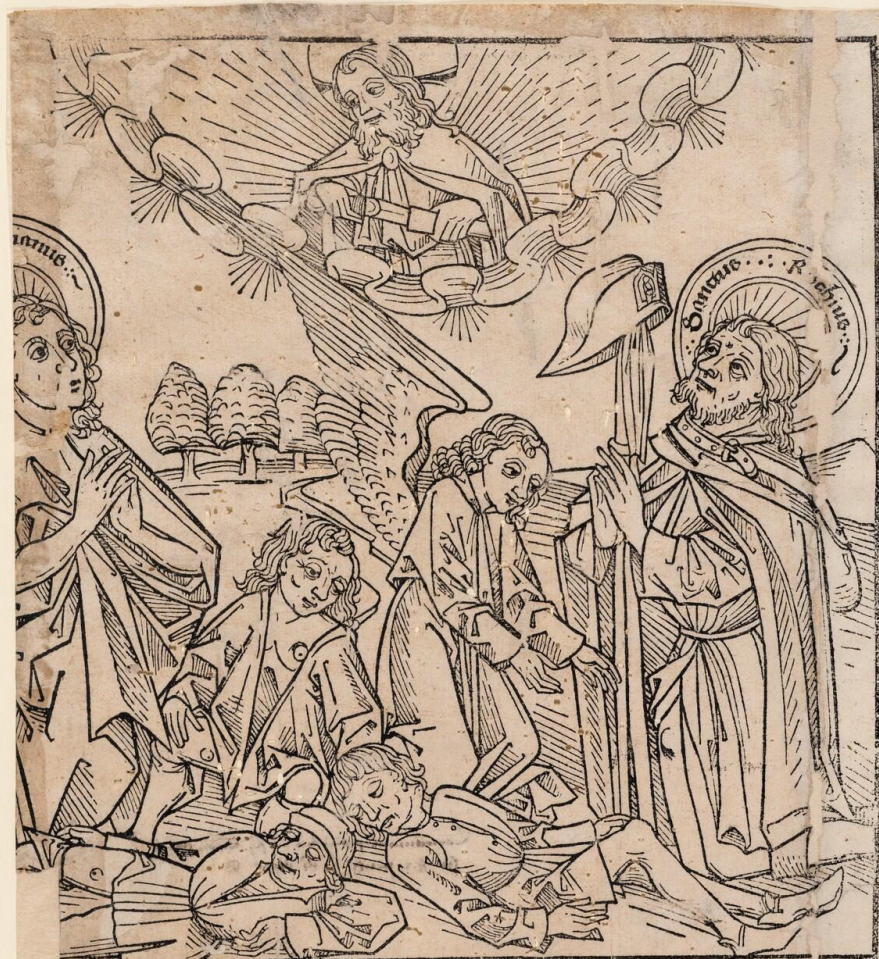
*Is this document related to the one in our base?  
Can we gather some data about entities from  
external sources, e.g., Wikidata?*





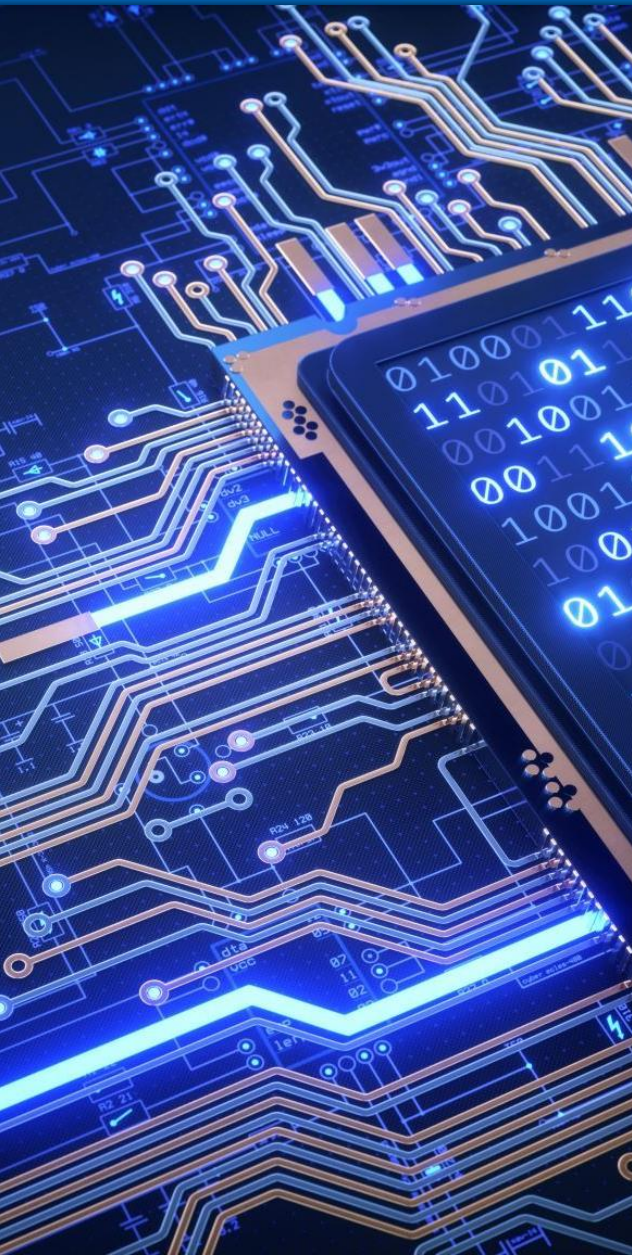
## Challenges

- 1) How to create a graph from tabular data?
  - 2) How can AI methods help to identify and correct errors (which are unavoidable)?
  - 3) How to prepare a **usable** hierarchy of identifiers to clearly distinguish units, documents and individual parts/pages?
  - 4) How can (semi-)automated modules enrich metadata, e.g., search and link other edits / comments / elaborations for each document?
  - 5) How to prepare OCR models for medieval German texts?
- It may be more tricky than contemporary texts...*



basi. Wie g. ist dem ge  
re. v. unser herren ihu xpi  
den. v. das preses der  
werde. bit für v. as du ha  
re. d. gelupf v. unser herren  
Allmechtiger ewiger  
edieren v. gebete des  
s. sant Sebastians vor de  
ffen der pe. n. l. entz. v. us ge  
nittern v. allest. Verleiche  
h. bitten. oder diß gebett  
an. d. achtig. g. l. chen. sprach  
vor dem gebrechen  
net. werden v. durch  
n. h. l. igen. v. us. vor al  
nd. an. g. l. fen. leybs. v. nd  
et. werden. Amen.

**D**u haluf beüchtig rochi die hat got  
großg. n. ad. er. n. g. so. n. g. e. off. d. er  
dir. gen. ad. ge. be. hat. re. tran. cte. ge. l. und  
z. em. a. che. die. bl. u. de. z. u. e. l. e. u. c. h. e. v. u. die. be.  
h. a. f. te. z. u. e. n. l. e. d. i. g. e. v. u. d. g. o. t. h. a. t. d. u. z. u. i. t.  
d. a. s. r. e. i. ch. s. e. m. s. h. u. m. l. i. s. c. h. e. n. v. a. t. e. r. e. s. E. w. i. g. e.  
g. o. t. i. ch. b. i. t. d. i. ch. a. l. s. d. u. d. e. m. e. h. a. l. i. g. e. b. e. t. t. e.  
i. g. e. r. r. o. c. h. i. o. g. n. a. d. e. r. h. a. n. g. t. b. a. s. t. a. l. s. e. z. w. a. l.  
l. e. t. z. u. d. e. h. a. l. i. g. e. s. t. e. r. t. e. i. n. d. e. m. e. n. a. m. e. s. v. o.  
d. e. m. v. e. r. g. i. f. t. e. g. e. s. c. h. w. e. r. z. w. i. s. c. h. e. s. e. m. e. b. e. n.  
e. n. v. u. s. m. i. s. s. f. d. e. n. t. o. d. l. e. i. d. e. n. t. d. u. r. c. h. a. n. w. e. l.  
u. n. g. d. e. m. e. s. h. a. l. i. g. e. n. n. a. m. e. n. s. i. n. t. d. e. d. i. a. n. s. t.  
d. e. m. s. h. a. l. i. g. e. n. e. n. g. e. l. e. s. i. n. h. a. s. t. g. e. h. a. l. e. t. v. u. d.  
f. i. r. o. s. e. m. v. o. u. n. s. a. m. z. g. e. f. e. r. t. g. e. w. e. s. e. n. b. i. s. t. v.  
l. e. i. c. h. v. u. s. g. e. n. e. d. i. g. t. l. i. c. h. a. l. l. e. d. i. e. d. e. m. e. n. g. o. t.  
l. i. c. h. e. n. n. a. m. e. n. a. n. w. e. l. f. f. e. n. d. u. r. c. h. d. a. s. g. e. b. e. t. t.  
v. i. r. v. d. i. e. n. e. s. a. n. t. r. o. c. h. i. o. s. o. f. o. l. l. i. c. h. e. v. e. r. g. i. f. t. e.  
b. r. e. s. t. e. v. u. d. d. e. e. w. i. g. e. t. o. d. e. r. l. o. s. t. w. e. r. d. e. n. t.  
d. u. r. c. h. v. u. s. e. r. n. h. e. r. e. n. i. h. u. x. p. i. n. A. m. e. n. z.



## Structure of the project: Platform Two

Research ideas & opportunities in the area of AI solutions:

- Automated discovery of semantic relationships between documents, persons, locations, etc.
- Social network reconstruction of authors, collectors, etc.
- Handwriting recognition of texts from different periods
- Beyond text: visual analysis and semantic annotations of documents containing pictures, musical score, etc.

## Concluding remarks & Research opportunities

- We are working with unique resources that are European cultural heritage
- Interesting challenges in knowledge engineering and opportunities to collaborate with the Jagiellonian Library already hosting large digital collections
- Opportunities for developing novel AI-based solutions in the area of digital libraries
- Interdisciplinary environment
- Starting platform for new specialized R&D projects with external research groups and companies



**European Treasure in the Jagiellonian Library.  
World Heritage in the former Prussian State Library.  
Digitization – Interdisciplinary research**



Thank you for your attention!