

Putting 600,000 books online

The large-scale digitisation partnership between
the Austrian National Library and Google

Max Kaiser

Head R&D, Austrian National Library

40th LIBER Annual Conference
Barcelona 29 June – 2 July, 2011

Austrian Books Online

www.onb.ac.at/ev/austrianbooksonline/

Digitisation

of the entire historical
book holdings of the
Austrian National Library

Largest Austrian Public Private Partnership in the cultural sector



Österreichische
Nationalbibliothek

Papyrusmuseum

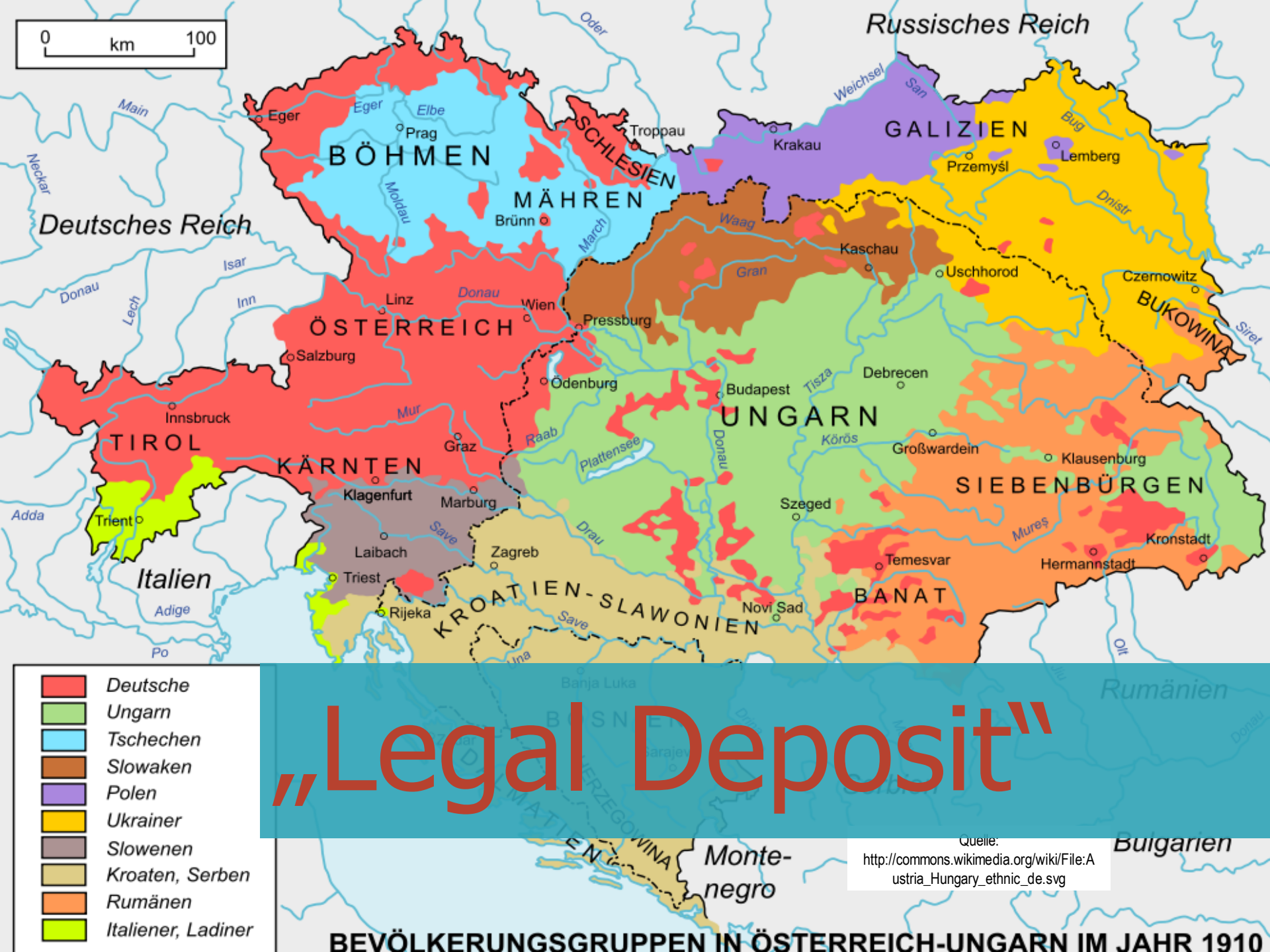
epheios museum
holzgerätekammer
sammlung alter musikinstrumente
khm

History back to the 14th century



One of the world's
most significant
historical collections





- Deutsche
- Ungarn
- Tschechen
- Slowaken
- Polen
- Ukrainer
- Slowenen
- Kroaten, Serben
- Rumänen
- Italiener, Ladiner

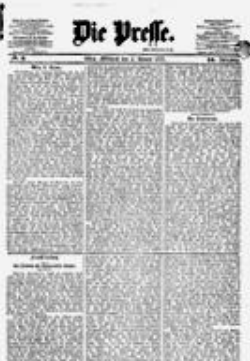
„Legal Deposit“

Quelle:
http://commons.wikimedia.org/wiki/File:Austria-Hungary_ethnic_de.svg

BEVÖLKERUNGSGRUPPEN IN ÖSTERREICH-UNGARN IM JAHR 1910

→ Digitisation

Tagesübersicht: 4. Januar 1871



K. K. Hof-Operntheater.
Wien, den 4. Januar 1871.
Prinzessin Wauwendenz.
Oper in drei Akten.
Musik von Robert Schumann.
Schauspiel von Eduard Scribe.
Ballett von Louis Perrot.
Besetzung:
Prinzessin Wauwendenz: Fräulein ...
König: Herr ...
Minister: Herr ...
etc.

K. K. Hof-Operntheater.
Wien, den 4. Januar 1871.
Caraculo-Mentore in Paris.
Oper in drei Akten.
Musik von Robert Schumann.
Schauspiel von Louis Perrot.
Ballett von Louis Perrot.
Besetzung:
Caraculo-Mentore: Herr ...
etc.

→ Digitisation

→ Online Publications

→ Web Archiving

6 years

600,000 volumes

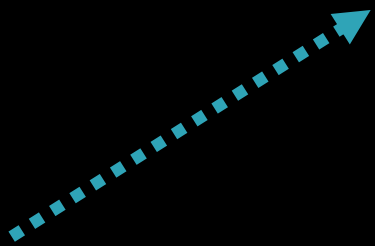
180 Mio pages



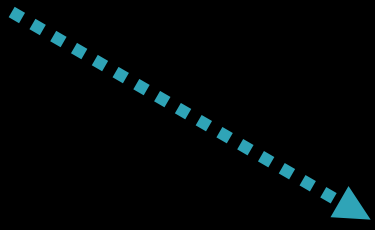
Google Books

Digital Library
Austrian National Library

Google Books



Partner Program



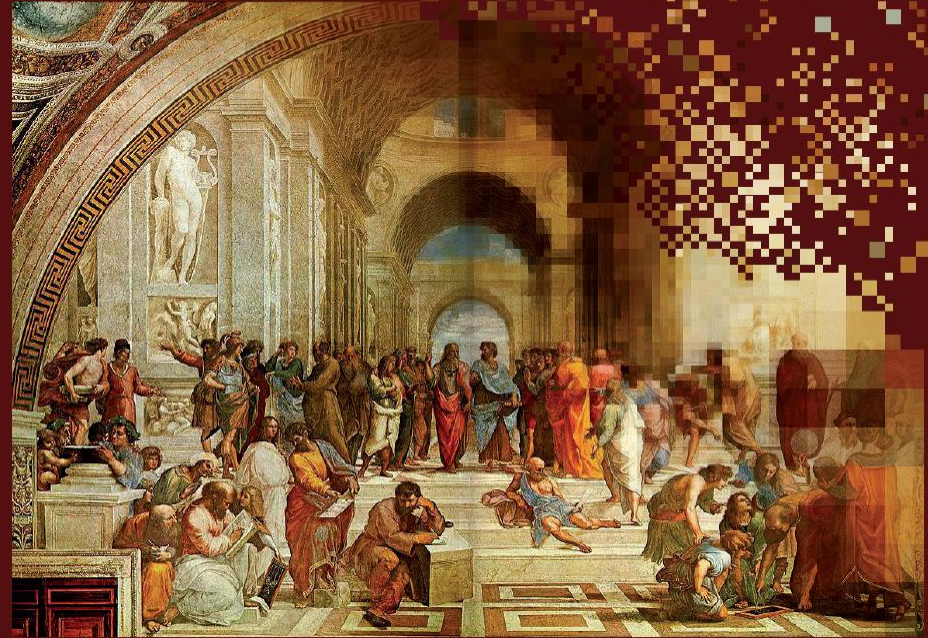
Library Program

13 Libraries in Europe

5 National Libraries

- Italy
- Austria
- The Netherlands
- Czech Republic
- Great Britain

THE NEW RENAISSANCE



http://ec.europa.eu/information_society/activities/digital_libraries/doc/reflection_group/final_report_%20cds.pdf



REPORT OF THE 'COMITÉ DES SAGES'

REFLECTION GROUP ON BRINGING EUROPE'S
CULTURAL HERITAGE ONLINE

„Stimulating the flow of private funds for the digitisation of cultural assets through equitable **public private partnerships** appears as a viable and sustainable way of tackling the pressing question of making Europe’s cultural wealth accessible online and preserving it for future generations.“

„The key question is not whether public-private partnerships for digitisation should be encouraged, but **how**, and **under which conditions**.“

Cornerstones:

- Respect for rights holders
 - Only public-domain works digitised
- Transparency
 - Very detailed FAQs online
- Access
 - Free-of-charge access worldwide

Cornerstones:

- Digital copies for library
 - Identical with copies used by Google
- Re-use
 - All files available for non-commercial use
 - Access via platforms like Europeana
 - Provision to research partners
- Non-exclusivity
 - Free to digitise material with other partners

Additional key elements:

- Selection of books by library
- Institute for Restoration involved
- Termination



Who is paying
for what?

Costs

- Full text-digitisation:
very expensive
- Report by
Collections Trust
for Comité des Sages

The Cost of Digitising Europe's Cultural Heritage

A Report for the Comité des Sages of the European Commission

Prepared by Nick Poole, the Collections Trust

November 2010

http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/annexes/digiti_report.pdf

70–100 Euro / book

→ Typical costs for projects with external service providers

Google:

- Transport
- Insurance
- Scanning
- OCR
- Image processing
- Quality control
- Google Books

Austrian National Library:

- Selection
- Internal logistics
- Conservational assessment
- Metadata
- Barcoding
- Data download and control
- Data storage & digital preservation
- Digital Library

Which books?

200.000 volumes

State Hall

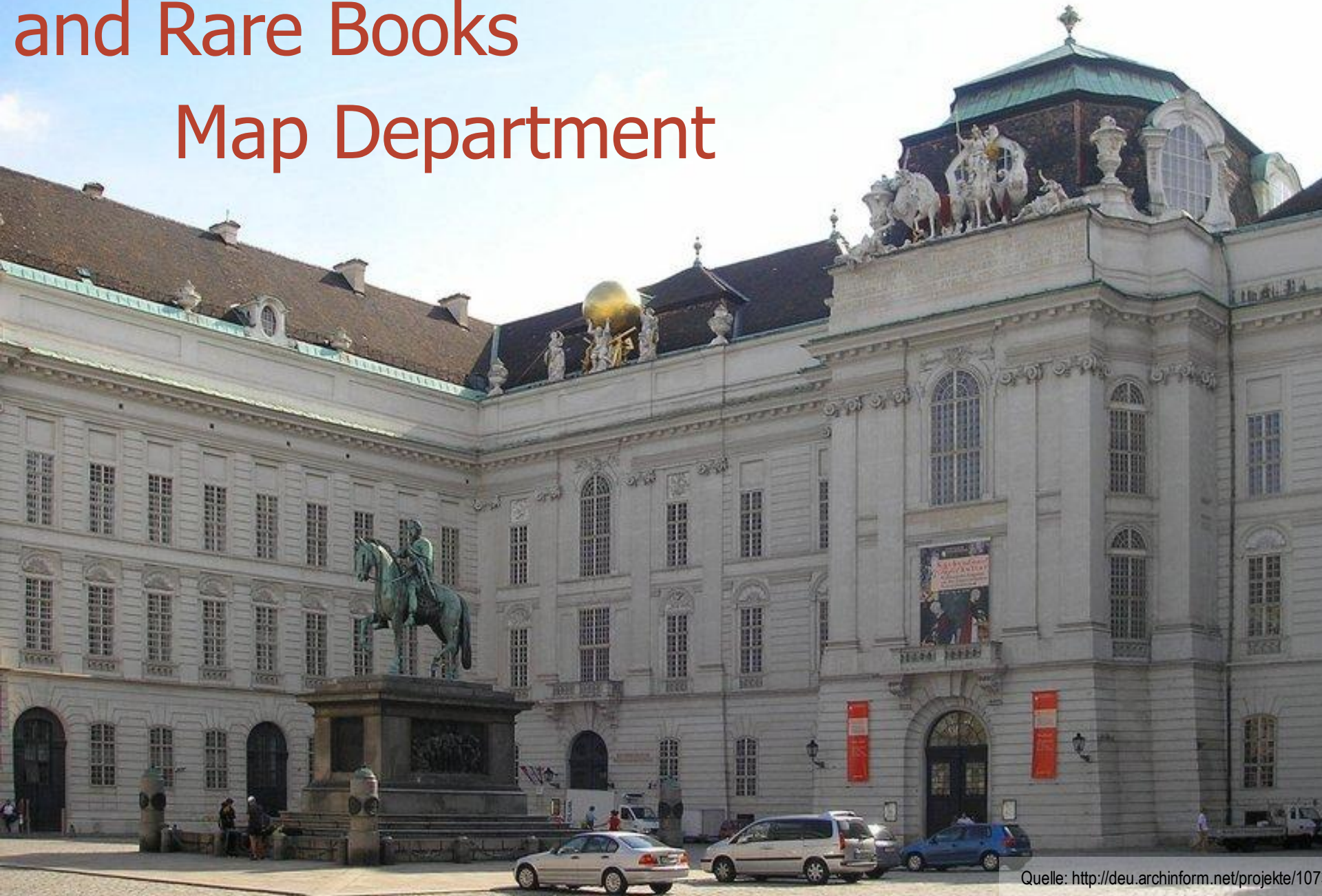




Entire historical
book holdings
16th –19th century

Department of Manuscripts and Rare Books

Map Department





Department of Music

Theatre Museum





Fidei Commiss Library

→ Conservation

→ Preservation



http://www.mediathek.at/akustische-chronik/popup/popup.php?document_id=1000115&zone_id=1000043&template_id=1000016&zone_name=IMAGE_ZONE1

7 Work Packages

- Book logistics
- Metadata / Catalogues
- Conservation / Restoration
- Data download / Quality control
- Access
- IT infrastructure
- Project management

Preparatory project

Mid - end 2010

- Integration with organisational processes
- Personnel resources
- Logistics workflows

Internal communication

- Change processes
- Re-evaluation of workflows
- Availability of internal resources

Consultation with other Google partners



70+ staff members

20+ exclusively for project

- Book logistics
- Metadata adaptation
- Cataloguing
- Conservation / restoration
- Quality control
- Software implementation
- Project management

End of 2010

Test shipment & Start operational project

Spring 2011

Start of digitisation



No individual selection ...

Format



Format



Condition





Preparation

Conservational evaluation



< 1%

excluded for
conservational reasons

Logistics in the State Hall



Logistics in the State Hall



Logistics in the State Hall



Challenges...



Challenges...



Challenges...



Challenges...



Challenges...



A man in a blue uniform is walking on a metal staircase in a warehouse. The warehouse has high shelves filled with boxes. The man is wearing a blue jacket and a yellow backpack. The staircase is made of metal and has a railing. The background shows a long aisle with shelves and a person in the distance. The lighting is bright, and the overall atmosphere is industrial.

Logistics in the „Aurum“ Depot



Logistics in the
„Aurum“ Depot

Preparation for Digitisation



A woman with short brown hair, wearing a blue lab coat over a black turtleneck, is standing in a library's manipulation area. She is holding a stack of several old, worn books and looking at them. The shelves around her are filled with more books, some with yellow and red labels. In the foreground, there is a wooden cart with a red handle. The text "Manipulation area ..." is overlaid in light blue on the left side of the image.

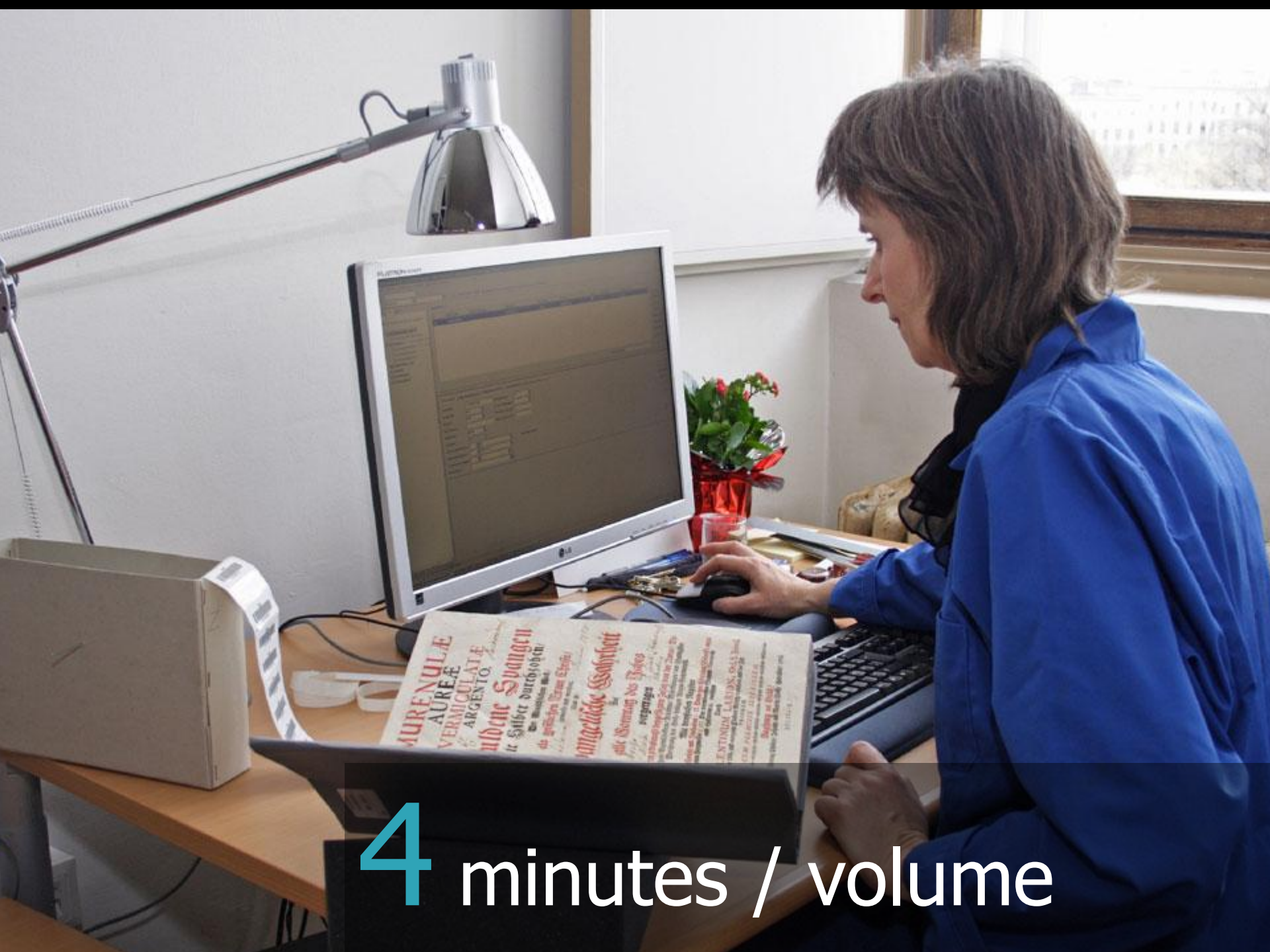
Manipulation area ...

Barcoding





Adaptation of metadata



4 minutes / volume

600.000 books

40.000 hours

5.128 working days

24,4 person years



Complex cases ...

Ad
Adel. 1
Adel. 4
Adel. 5
Adel. 7
Adel. 13

Osterreichische
Nationalbibliothek
227.615-D
Alt-

Adligata ...

QUAE TU SAUERRAY ANUPER DOCUMENTA DEDISTI.
 Dedit voluitque ac Latini vultusque per orbem.
 Ista ferunt sane Tuis nunc vultusque per orbem.
 abim rutilo. Doctoris pulchro abire iubenti.
 atq; caput summo Tibi latens ac vultus iubenti.
 Hinc rursus grati Tibi decantantur honores.
 placuisse ac bonis Te carmina multa in usque
 qua summas quot ei nactus gratantur honores.
 Hec inter verba mea sunt quoq; pauca verba.
 quae ego conor Tibi deo meranda Patrone.
 bisse novis applaudere honoribus atq; precari
 praeparatis. Ut confensio hocce cacumine honoris
 summo, In veras felices Nestoris annos.
 complures fructus signas es in agmine sacro.
 atq; coronari tandem excellenti corona
 vita quam Tibi des Dominus post secula multas.

Pauci hic & Excellentissimo Do. Doctori SAUERRAY,
 suo quondam Praeceptori fidalissimo, ex debitis ob-
 servantiis gratulatur Athenis Noticia

Theod. Anshelmus Gleiß/
 Philof. stud.



EYDHMIAI ILLUSTR. GYMNASII
 COBURGENSIS

REGIAE TIBI praesentia pueri Doctoreque Doctori
 Aquo DEO Doctus fingula puncta tollit.
 Te illustri Cathedra & Gerae pulchrae munitur.
 Quodsi Docturae rem inchoantur iuris
 san. Civ. in. Civ. Gerae & Gersperis hincque vultus
 Jobann Augustus Schreyer.

MAGIS magis dicit. Dum stultas vultus, quae non
 alim magnam & cultum soligit ingenuum.
 Ardua sunt, magno ut solus illi labor parali.
 sic olim TE, vultus dicit, magis Maturus.
 et Reverende Dominus, Georgius vry, Compagn. Theologiae h-
 nore exaltatus episcopus
 Johann Lucas Ahern, D. in M. Gymn. Cob. Prof. P. & Physicae Prof.
 nore exaltatus episcopus

Was ist Gefehsamkeit / bey den / der nicht gemeint
 An Log zu legen Sie? Ein Licht das gar nicht scheint /
 Ein tieff, verborgener Schatz / der in der Erden liegt /
 Ein Blätter voller Baum / der sonder Feucht verbleit.
 Wann nun / HÖCHSTWÄRTIGSTEN / Er sogetam ist zu legen
 Uns Loge / Licht den Schatz / den Er durch Gottes Segen
 Schon lang gesammelt hat / nebst großer Müß und Fleiß
 So reichet Palacom / Ihn jetzt den Doctors - Preiß.
 Drum Gratulire ich / der Himmel wolle geben /
 Daß in dem Stande Er viel Jahre mög leben /
 Auch unser Kirch und Schul verplure seine Lehr /
 Und bey der Jugend noch den Wachstumb fere mit
 Johans Ernst Dohz / Jur. Licent. von Feil.

Illa dies haec est & que Tu celebraveris amica.
 Qua Tibi gratantur toto de pectore amici.
 Cum nunc, ecce, Tibi titulos & grandia praemia
 Nomina, quae aeternae famae sub imagine fulgent.
 Et mea, quantumvis tenuis, nunc Musa sereret,
 Nec titulusque Tuis gratando vota referret,
 Quos virtus, pietas, doctrina, modestia, candor,
 Sedulitas, Studium Tibi pulchre promeruerit?
 Ita age! Magne veni Mularum Doctor ad istam,
 Summas & celebres cape, qui Tibi dantur honores.
 Gratulor atque precor, bene sint Tibi prospera quaeque,
 Quae capis, & maneat tua fama perennis in orbe.

Es bricht der Jugend Glang / der Weisheit Krafft und S
 In vollen Flammen aus / wann in der Seel und S
 Sich vor des Nächsten wohl erregen liebes Werke /
 Wann uns das Glück und Recht das Houbt zu heben
 Ein unterirdisch Feuer / wo es kan Lust bekommen /
 Durcht Stein und Eisen aus / und steigt Himmel an /
 Hat wohl des Menschen Macht sich jemal unternommen /
 Zu dämpffen diese Flamm? Nein / Nein / kein Mensch es kan.
 Es reißt die Erden auff / und bricht mit Knall und Krachen /
 Ein sich in tieffsten Schlund der Erd versteckter Wind /
 Er kan / was Felsenhart / zu Stück und Splutter machen /
 Und zeigt die meiste Stärck / wann Widerstand sich find.
 So auch HÖCHSTWÄRTIGSTEN / erhebt sich Seine Tugend /
 Die sein Gemüth gefast / und seine Seele ziert.

227645-9
 Ad. 70

Die

Adligata ...

Adl. 41	Adl. 53
Adl. 42	Adl. 54
Adl. 43	Adl. 55
Adl. 44	Adl. 56
Adl. 45	Adl. 57
Adl. 46	Adl. 58
Adl. 47	Adl. 59
Adl. 48	Adl. 60
Adl. 49	Adl. 61
Adl. 50	Adl. 62
Adl. 51	Adl. 63
Adl. 52	

Adl. 64	Adl. 76
Adl. 65	Adl. 77
Adl. 66	Adl. 78
Adl. 67	Adl. 79
Adl. 68	Adl. 80
Adl. 69	Adl. 81
Adl. 70	Adl. 82
Adl. 71	Adl. 83
Adl. 72	Adl. 84
Adl. 73	Adl. 85
Adl. 74	Adl. 86
Adl. 75	Adl. 87

Adligata ...



„Slim“ volumes ...



Special collections ...



Conservational protection



Conservational protection



Conservational protection



Cataloguing of the Fidei Commiss Library



Cataloguing of the Fidei Commiss Library



Ready for Digitisation ...

Digitisation

- Scanning Center in Germany
- Procedures agreed
- Austrian Federal Office for Monuments involved
- Each volume checked after return
- Books unavailable to users for ~ 3 months

Book Logistics> Digitisation



Data Download

Quality Control

Storage

Access

ADOCO
(Austrian Books Online
Download & Control)

30 Mio.
digitised items / year

143.000
digitised items / day

Quality control

- Automated jobs
- Representative samples
- IT assisted discovery of error clusters
- Error candidates checked manually

- Crowdsourcing?

Data Management & Access

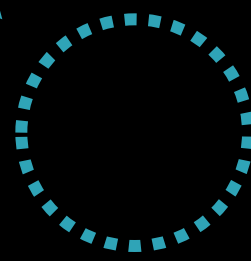
- Data storage: inhouse
- JPEG-2000 master files stored redundantly
- Access copies generated on-the-fly
- URN resolver for permanent identification



europaena
think culture

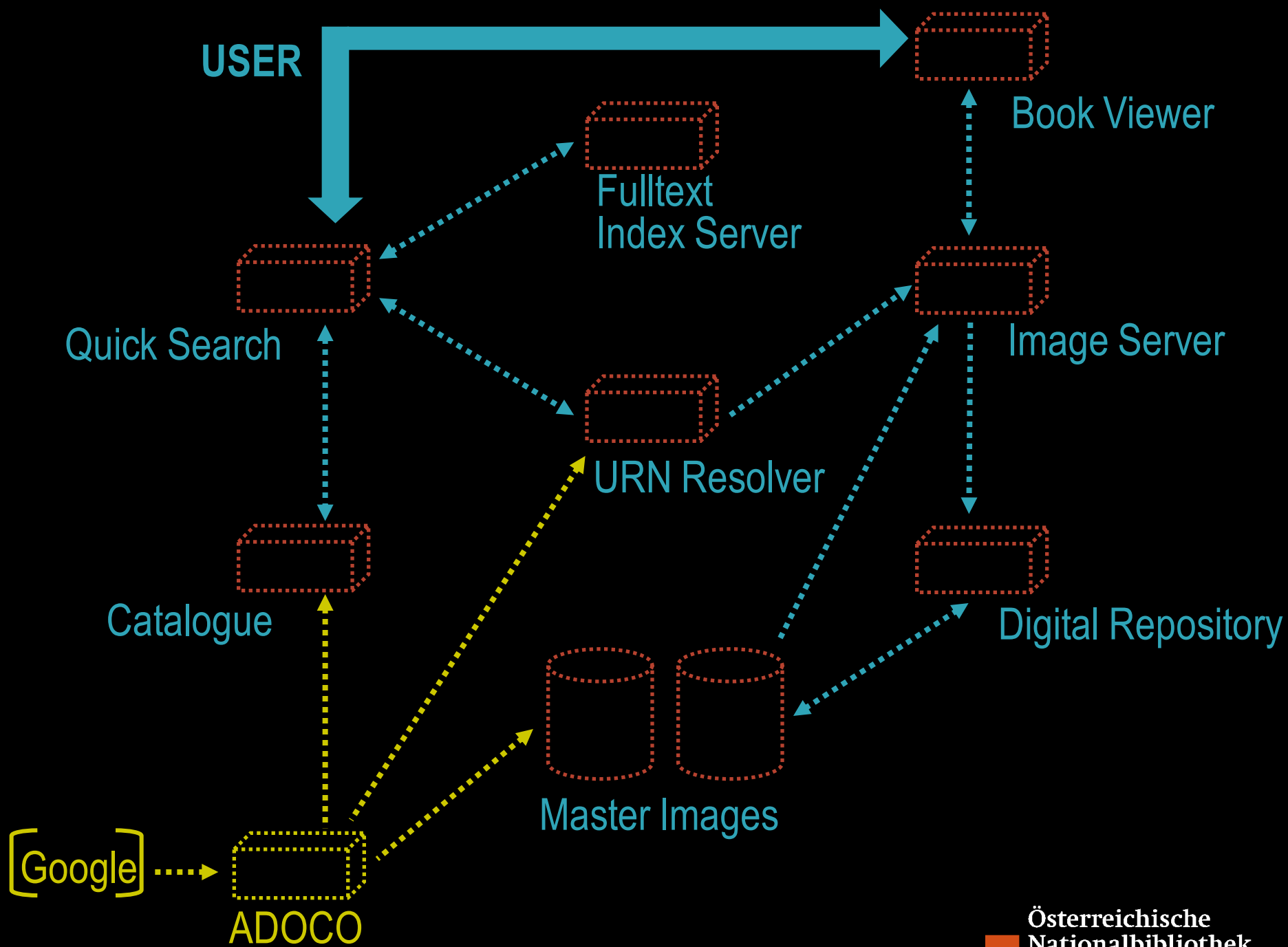
Catalogue /
“Quick Search”

Full-Text Search



Book Viewer

[Mobile Apps]



Outlook

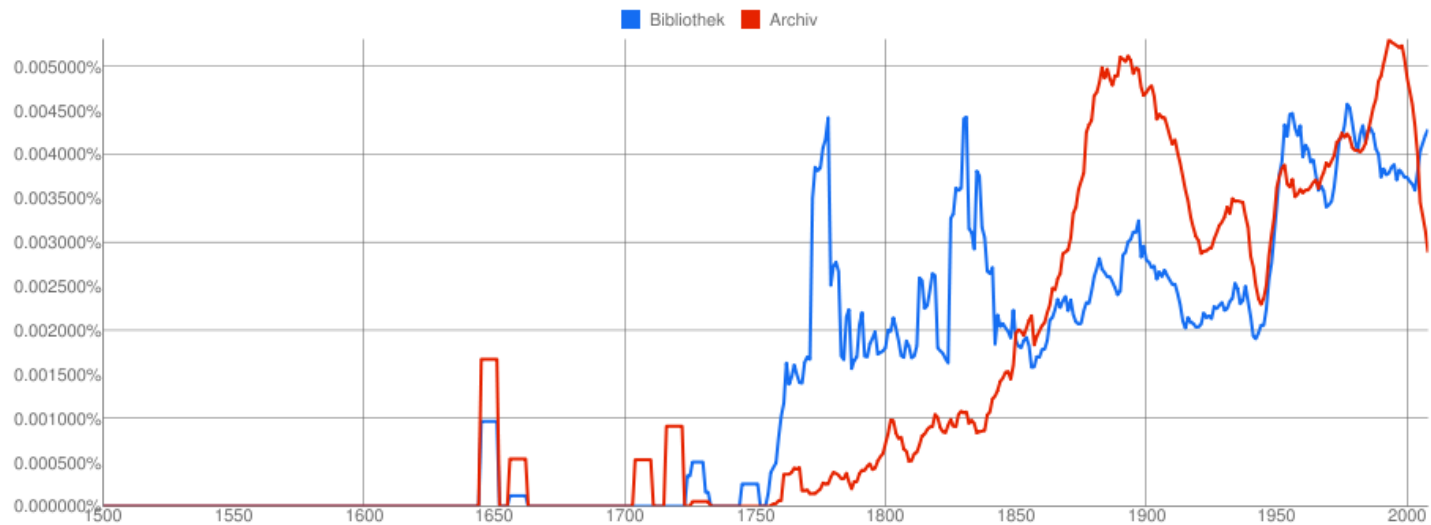
- Full-Text: new possibilities for research
- Data enrichment
- Named Entity Recognition
- Linked Data
- New data centric research in the Humanities & Social Sciences



<http://www.diggingintodata.org/>

Graph these **case-sensitive** comma-separated phrases:

between and from the corpus with smoothing of .



0
 0

Search in Google Books:

1500 - 1782	1783 - 1956	1957 - 1975	1976 - 1991	1992 - 2008	Bibliothek (German)
1500 - 1833	1834 - 1887	1888 - 1901	1902 - 1992	1993 - 2008	Archiv (German)

Run your own experiment! Raw data is available for download [here](#).

★★★★★
 0 Reviews
 Write review
 About this book

Search in this book Go

Add to My Library

 **Google eBook** New!
 Buy once. Read anywhere. [Learn more](#)

Free
 Better for larger screens. [?](#)
[Get it now](#)
[View sample](#)

[Read on your device](#)

- Get this book**
[AbeBooks](#)
[GOOP](#)
[On Demand Books](#)
[Amazon](#)
[Find in a library](#)
 All sellers >

Related books



All related books >

Huy! und Pfuy!
Der Welt.
Huy/ Oder Aufrischung
 Zu
 allen schönen Tugenden:
Pfuy Oder Abschreckung
 Von
 allen schändlichen Lastern:
 Durch unterschiedliche sittliche Concept,
 Historien/ und Fabeln vorgestellt.
 Worinnen
Der Hoet / Prediger / und waserley
 Standes- Personen für ihren Gram
 etwas finden können;
 Durch
R. P. ABRAHAM a S. CLARA,
 Augustiner Barfüßer Ordens / Provinciae
 Definitorum, und Kaiserl. Prediger x. x.
 Mit Römischer Kayserlicher Majestät aller gnädigsten Freyheit/
 Mit Kupfern gezieret und verlegt
 Durch
Christoph Weigel/ Kupferstechern und Kunsthandlern
 in Nürnberg.

Würzburg/
 Gedruckt bey Martin Franz Herzen.
 1707.



★★★★★
0 Reviews
[Write review](#)
[About this book](#)

[Add to My Library ▾](#)

 **Google eBook** New!
Buy once. Read anywhere. [Learn more](#)

Free

[Read on your device](#)

Get this book
[AbeBooks](#)
[QOOP](#)
[On Demand Books](#)
[Amazon](#)
[Find in a library](#)
[All sellers >](#)

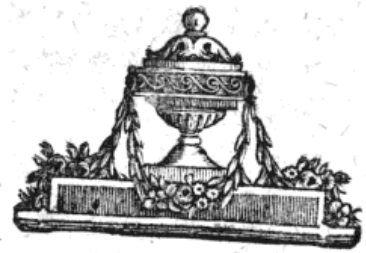
Related books
  

[All related books >](#)

Rede
wider das
Lesen
böser Bücher.



Herausgegeben
von einem Priester aus dem Lambertischen
Kollegio der Adlichen zu Trier.
Cum approbatione Ordinariatus Trevirensis.



Zum zweytenmal aufgelegt im Jahre 1784.

297.284.-A

More information

www.onb.ac.at/ev/austrianbooksonline

www.onb.ac.at/ev/austrianbooksonline/faq.htm

twitter.com/abooksonline

Thank you!

max.kaiser@onb.ac.at

www.onb.ac.at

www.linkedin.com/in/maxkaiser

twitter.com/maxkaiser