

Austrian Books Online

Google Books based mass digitisation

Stefan Majewski

OPF Hackathon 2.12. - 4.12.2013

Austrian National Library, Vienna

Overview

→ The project

→ How the data is acquired, from carrying the book to storing the files.

→ The delights and perils of mass digitisation

→ Some challenges

→ How to work with the data?

→ Data organisation

Austrian Books Online

The Project

Key Facts

- Scope: 600,000
 - 200 Mio Pages
- Progress: 180,000 -> 5,500/3weeks
- Workforce: 20+ FTE -> 60+ P
- Areas
 - Logistics
 - Metadata
 - Conservation
 - Download & QA
 - Online Presentation
 - Storage
 - PM

Material

→ legal deposit >> wide variety of material

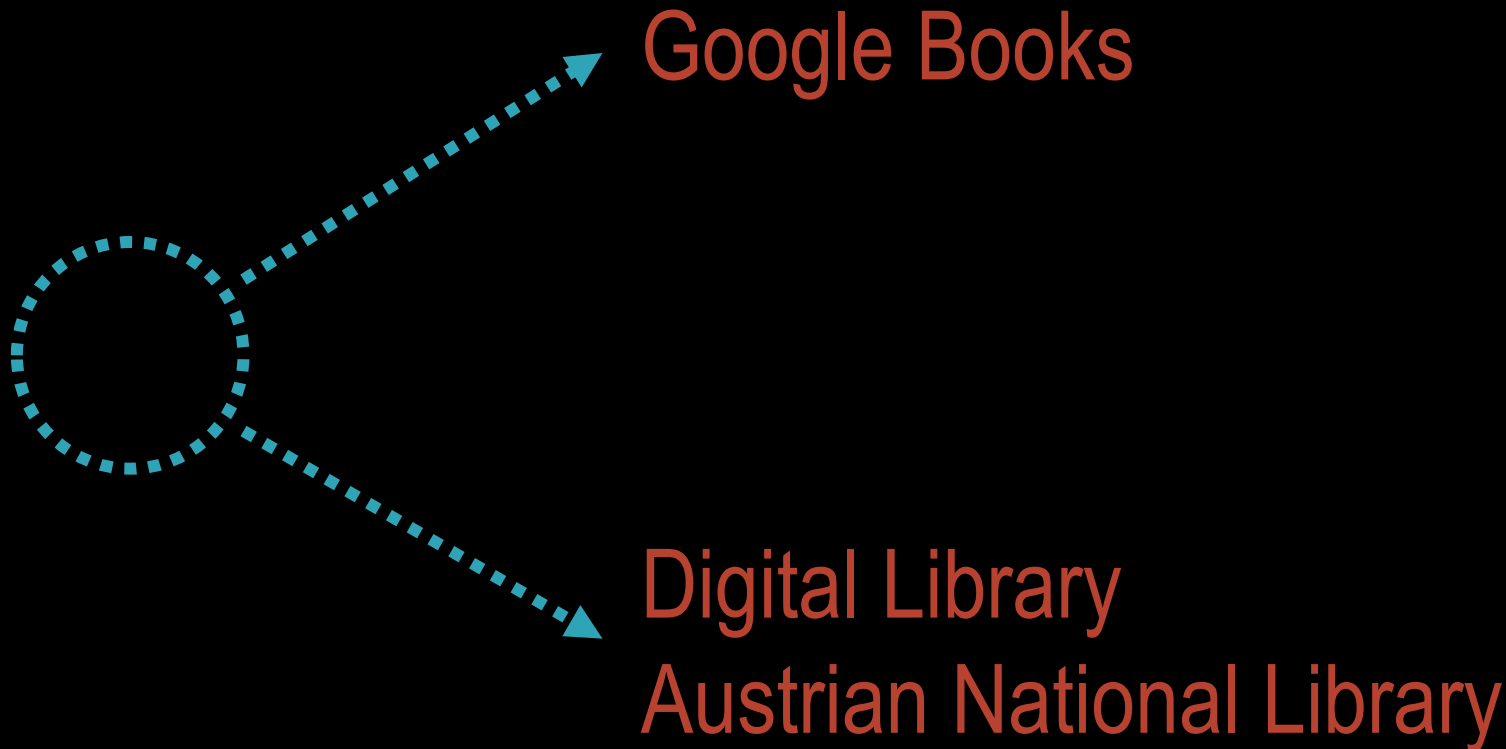
→ from:

16th century



19th 2nd half of
century

Public Access



13 Libraries in Europe

5 National Libraries

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

> 20 Mio. books

> 50% non-English

~ 75% from libraries

~ 2 Mio. books from European libraries

> 3 Mio. books public domain

digitisation

of the entire historical
book holdings of the
Austrian National Library

16th to 19th century

70+ staff members

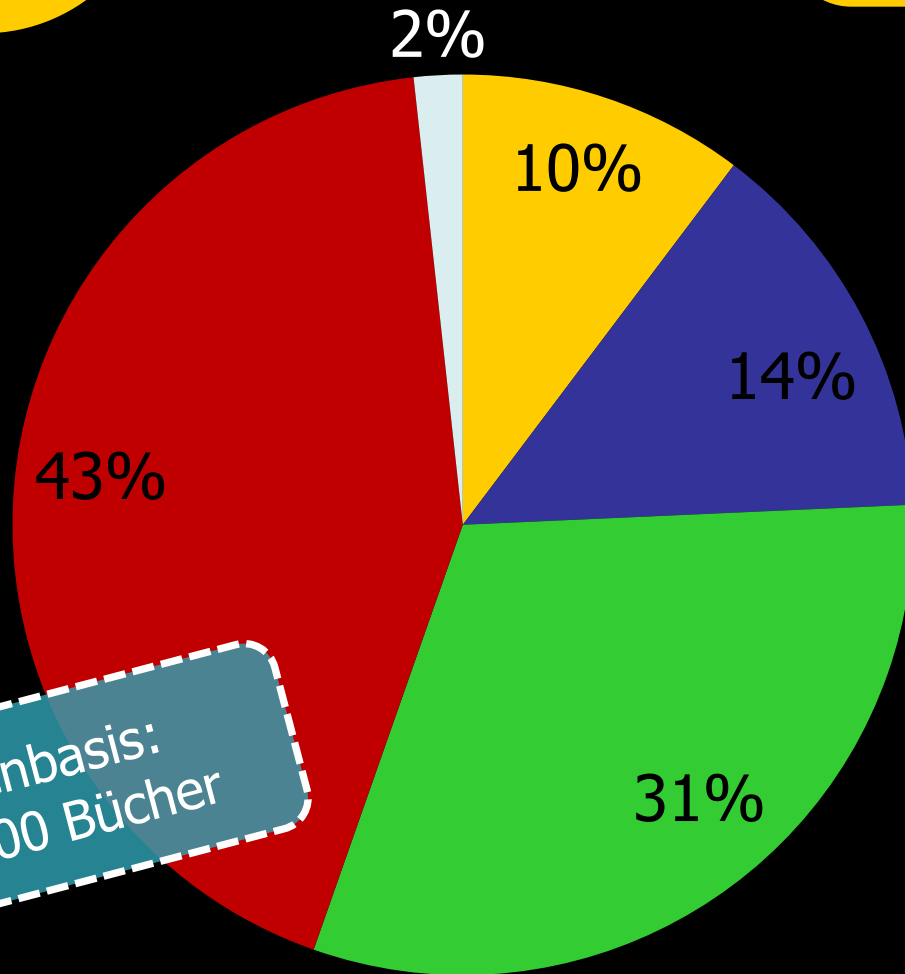
20+ exclusively for project

- book logistics
- metadata adaptation
- cataloguing
- conservation / restoration
- quality control
- software implementation
- project management

48,8 person years

Austrian Books
Online

Jahrhunderte ...



Datenbasis:
170.000 Bücher

■ 16. Jh.

■ 17. Jh.

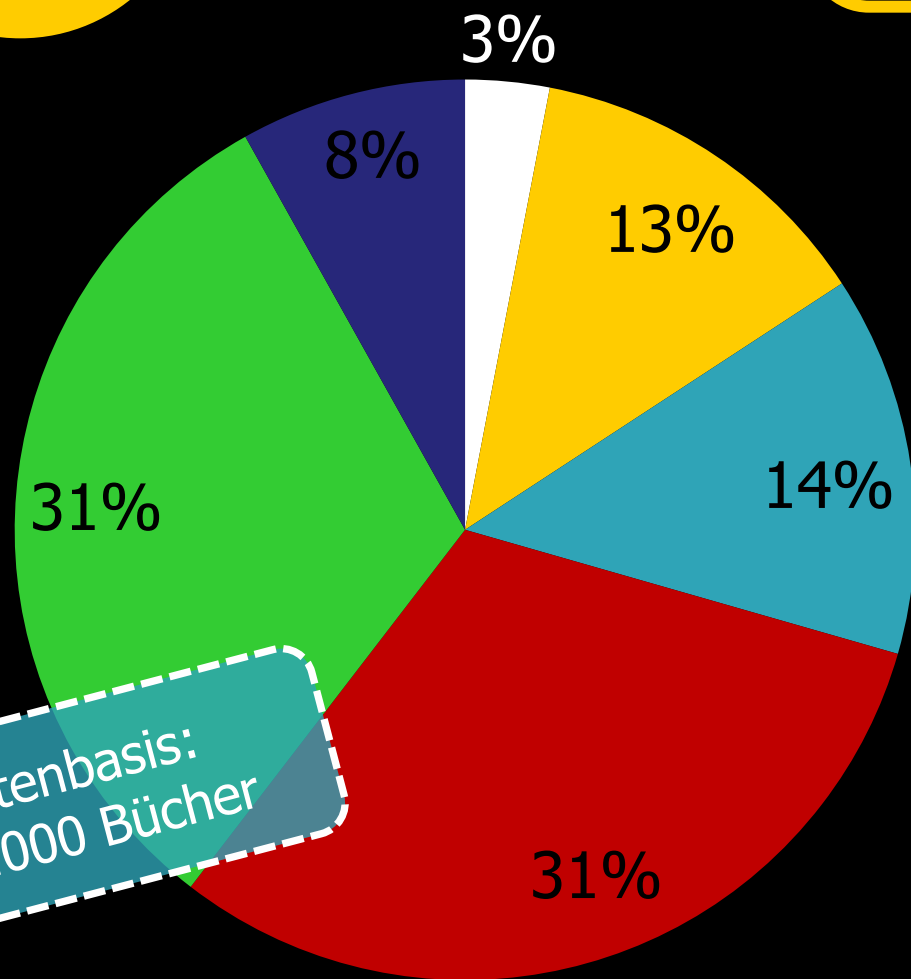
■ 18. Jh.

■ 19. Jh.

■ no year

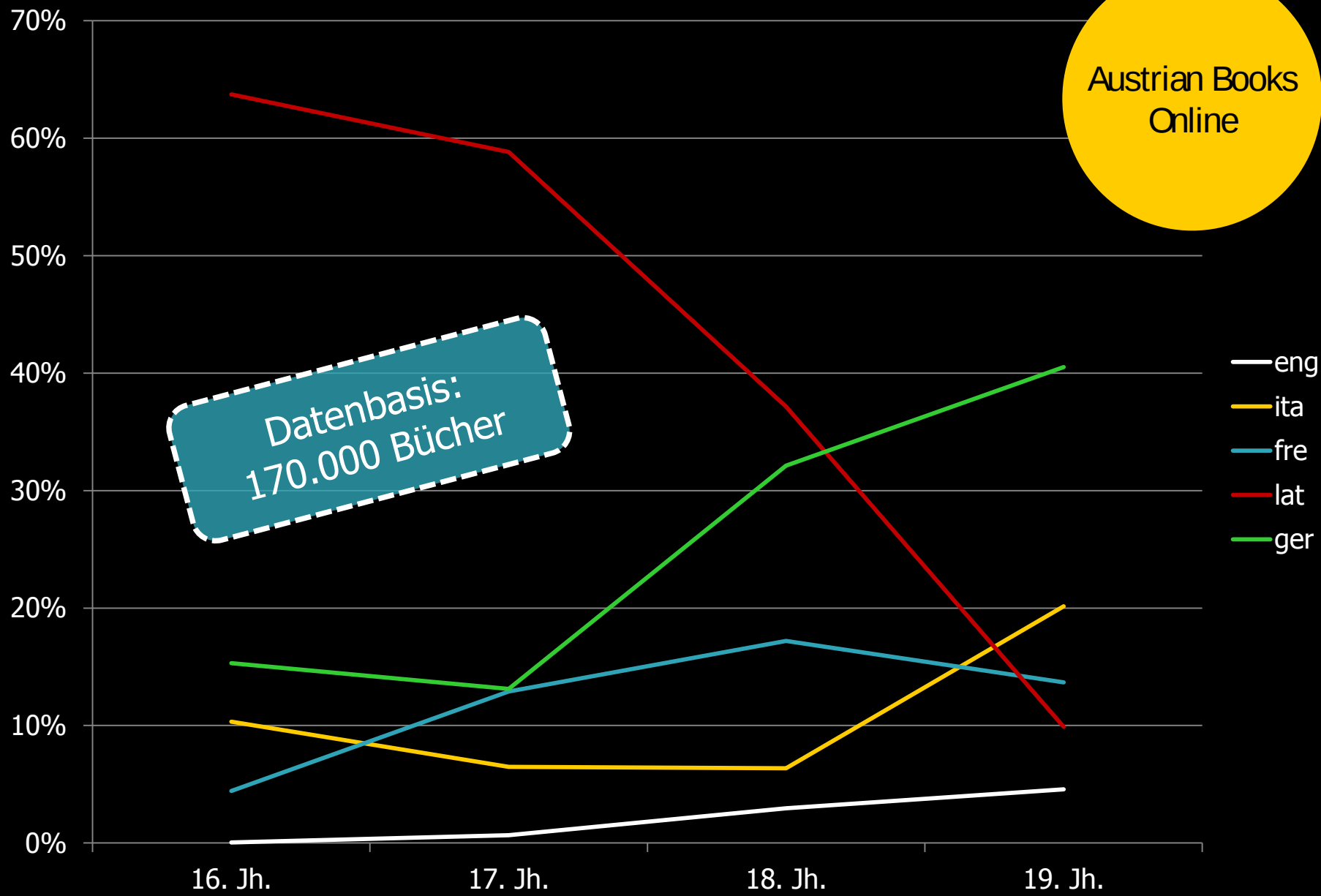
Austrian Books
Online

Sprachen ...

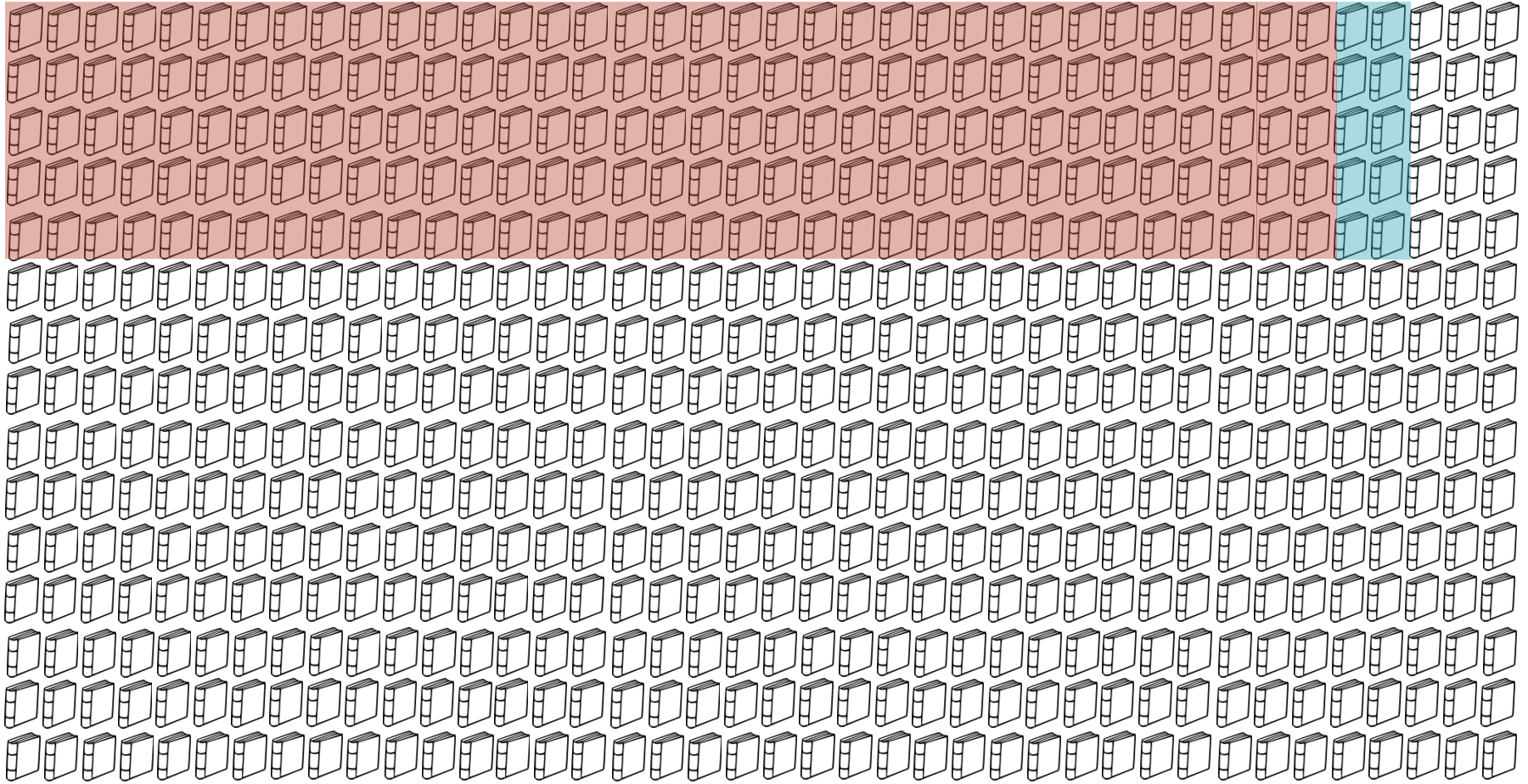


Datenbasis:
170.000 Bücher

eng
ita
fre
lat
ger
others



Ende 2013



~185.000 Bände digitalisiert

ÖNB Buch-Viewer

Volltextsuche

52+ Millionen Seiten

1+ Milliarde unterschiedliche Terme

[Alle Bestände](#)[Druckschriften](#)[Digitale Ressourcen](#)[Bilder und Grafiken](#)[Sonderbestände](#)[Historische Volltexte](#)

gratz steyermark

[Suche](#)[Erweiterte Suche](#)

rss

[Seite dem e-Shelf
hinzufügen](#)**Ergebnisse verbessern****Thema**[Tageszeitung \(3.053\)](#)[Deutsch \(337\)](#)[Zeitung \(242\)](#)[Wochenzeitung \(118\)](#)[Graz \(99\)](#)[Mehr Optionen ▾](#)**Autor/in**[Jansekowitsch, Max \(83\)](#)[Lauer, Joseph Cölestin \(58\)](#)[Schickh, Johann \(25\)](#)[Gött, Janos \(10\)](#)[Beck, Christian-Daniel \(8\)](#)[Mehr Optionen ▾](#)**Erscheinungsjahr /****Datierung**[Vor 1786 \(417\)](#)[1786 Bis 1824 \(6.653\)](#)[1825 Bis 1863 \(14.729\)](#)[1864 Bis 1903 \(6.226\)](#)[Nach 1903 \(2.423\)](#)[Mehr Optionen ▾](#)**Medium**[Zeitschrift/Zeitung \(3.071\)](#)[Buch \(729\)](#)[Mehr Optionen ▾](#)**Ergebnisse 1 – 10 von 3.800 für Historische
Volltexte**[1](#) [2](#) [3](#) [4](#) [5](#) [→](#)

sortiert nach:

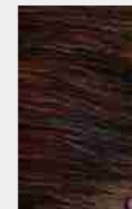
[Erscheinungsjahr ▾](#)

1



Buch

**Des loblichen Fürstenthumbs Steyr Erbhuldigung, inn dem
Fünfftzehnhundert vnnd zwaintzigsten, Auch nochuolgend des
ainundzwainzigsten Jars bescheen, sambt handlung der Landtäg
des selben mals gehalten, mit ettlichen ihren Freyhaitten,
Landshanduest vnnd beröff der ordnung in gemainen
handtierungen**

[Hofmann, Hans](#)[Augsburg : Necker ;1523](#)[● Online-Zugriff frei](#)[Online Ressource](#)[Details](#)

2



Buch

**Der durchleuchtigsten Erz-Herzogen zu Österreich Leben,
Regierung und Groß-Thaten. 1. Von dem aller-preiswürdigsten
Urheber dieses ... Erzhauses, Rudolpho, Grafen von Habsburg, so
wol aus diesem Haus ... erstem Römischen Kayser an, biß in die
Regierung ... Leopoldi, und ... Josephi. Aus bewährten
Geschichtschriften verfasst ... von J. C. B.**

[Beer, Johann Christoph ;J. C. B.](#)[Nürnberg : Endter ;1695](#)[● Online-Zugriff frei](#)[Online Ressource](#)[Details](#)

3



**Marianisches Jahr-Buch, in welchem gehandelt wird was betrifft
das Wunder-Gnaden-Bildlein Mariae zu Lankowitz in Unter-
Steyrmark gelegen (etc.)**

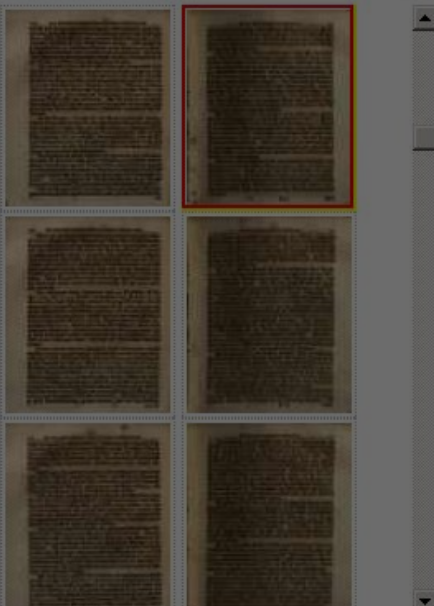
[Hietling, Conrad](#)[Wien : Hofbuchdruckerey ;1720](#)

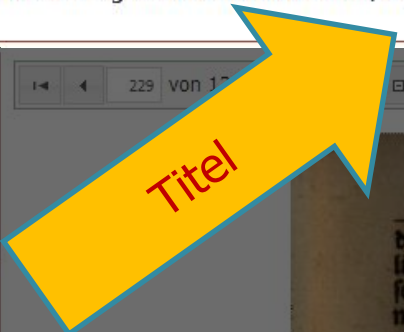
Information

Der durchleuchtigsten Erz-Herzogen zu Österreich Leben, Regierung und Groß-Thaten. 1. Von dem aller-preiswürdigsten Urheber dieses ... Erzhauses, Rudolpho, Grafen von Habsburg, so wol aus diesem Haus ... erstem Römischen Kayser an, biß in die Regierung ... Leopoldi, und ... Josephi. Aus bewährten Geschichtsschriften verfasst ... von J. C. B. Beer, Johann Christoph; J. C. B. Nürnberg: Endter; 1695
Zitierlink:
<http://data.onb.ac.at/ABO/%2BZ180063809>

Suche

"gratz" "steyermark"  

Übersicht




Leben, Regierung und Absterben. 171

der vor die Hand nahm / wurde Er von Selbigem mit 300. Geistlichen Pfründen beschenkt / um mit selbigem / wann sie ledig würden / seine Günstlinge zu erfreuen. Er selbst aber wurde samt seinem Comitatz nicht allein zu Rom / sondern auch im ganzen Kirchen-Staat allenthalben frey gehalten.

Wie Er An. 1469. wieder zurucke came / fand Er / wie jenes mals / seine Erblande in schlechtem Zustande / indeme etliche Landherren / unter denen Andreas Baumkircher / und beede Hannsen von Pefing und Stubenberg die Vornehmsten waren / wegen einiger ausschender Gold-Gelder / die sie nicht erlangen konnten / in Nieder-Steyermark feindlich einfielen / fünf an den Ungarischen Grenzen geieuge Schwärmer eroberten / mit Mannschafft besetzten / und sich damit unter Königs Matthiae Schutz begaben / welcher ohne dem mit Kayser Friderichen nicht zum besten stunde / weil ihme die / wegen des mit seinem Schwere / dem Böhmischen König / führenden Krieges / versprochene Gelder nicht gertichet wurden / wodurch endlich ein weitläufftiger Krieg entstanden / wie wie etwas besser unten vernehmen werden.

Als dieser Steyerische Krieg schon in die 5ten Jahre gedauert / und der Kayser dem gemeldten Baumkircher in die 700. Mann zu schanden gemacht / dieser hingegen dem Kayser bey 400. Mann theils abgefangen / theils niedergemacht hatte / auch Baumkircher wol vermerte / daß Er in die Länge den Handel nicht würde fortführen können / langte er um sicher Gelait an / um sich bey dem Kayser zu verantworten. Sein Begehren wurde ihme hierauf in so weit zugestanden / daß er von einer Vesperglocke zur andern sicher Gelait haben / und er sich am S. Georgen Tage zu Grätz einfinden sollte / wie auch geschah.

Nachdem er nun verhört wurde / und allerhand Ausflüchte / sein Thun zu rechtfertigen / auf die Bahn brachte / ruckte das Ende des Termins herben / welchen er von Kayser Friderichen zu verlängern bate / der sich mit seinen Rächten deswegen zu bereben vorschützte / und so lange damit verzogte / bis die Zeit des Gelaits zu verfließen begunnte / dannhero Baumkircher vom Schloß herunter liefe / sich auf sein Pferd schwangete / und damit zum Thor hinaus zu rennen trachtete.

D ij Weil

Information

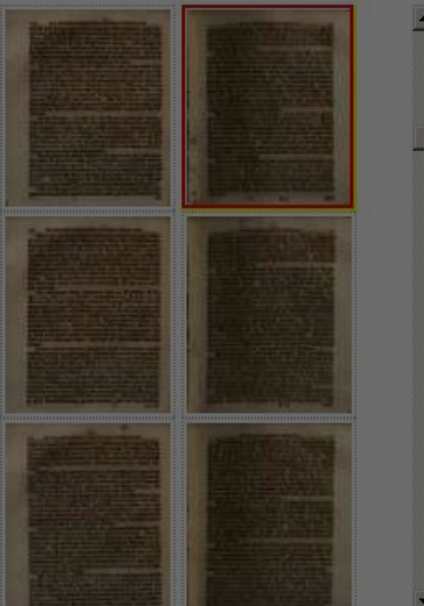
Der durchleuchtigsten Erz-Herzogen zu Österreich Leben, Regierung und Groß-Thaten. 1. Von dem aller-preiswürdigsten Urheber dieses ... Erzhauses, Rudolpho, Grafen von Habsburg, so wol aus diesem Haus ... erstem Römischen Kayser an, biß in die Regierung ... Leopoldi, und ... Josephi. Aus bewährten Geschichtsschriften verfasst ... von J. C. B.

Beer, Johann Christoph; J. C. B.
Nürnberg: Endter; 1695
Zitierlink:
<http://data.onb.ac.at/ABO/%2BZ180063809>

Suche

"gratz" "steyermark"

Übersicht



229 von 1290

Bild 1 von 7 mit Treffern

Ansicht

Leben/Regierung und Absterben. 171

die Hand nahm/ wurde Er von Selbigem mit 300. Geist-
fränden beschenkt/ um mit selbigen/ wann sie ledig würden/
anflinge zu erfreuen. Er selbst aber wurde samt seinem Co-
cht allein zu Rom/ sondern auch im ganzen Kirchen-Staat
allenthalben frey gehalten.

Wie Er An. 1469. wieder zurucke came/ fand Er/ wie jenes-
mals/ seine Erblande in schlechtem Zustande/ indeme etliche Land-
herren/ unter denen Andreas Baumkircher/ und beede Hannsen
von Pefing und Stubenberg die Bornehmsten waren/ wegen ein-
ger ausschender Gold-Gelder/ die sie nicht erlangen konnten/ in
Nieder-Steyermark feindlich einfielen/ fünf an den Ungarischen
Grenzen geieuge Schwärmer eroberten/ mit Mannschafft besetzten/
und sich damit unter Königs Matthiae Schutz begaben/ welcher
ohne dem mit Kayser Friderichen nicht zum besten stunde/ weil
ihme die/ wegen des mit seinem Schwere/ dem Böhmischem Kö-
nig/ führenden Krieges/ versprochene Gelder nicht gertichet wur-
den/ wodurch endlich ein weitläufftiger Krieg entstanden/ wie wie
etwas besser unten vernehmen werden.

Als dieser Steyerische Krieg schon in die 5ten Jahre gewäh-
ret/ und der Kayser dem gemeldten Baumkircher in die 700. Mann
zu schanden gemacht/ dieser hingegen dem Kayser bey 400. Mann
theils abgefangen/ theils niedergemacht hatte/ auch Baumkircher
wol vermerkte/ daß Er in die Länge den Handel nicht würde fort-
führen können/ langte er um sicher Gelait an/ um sich bey dem
Kayser zu verantworten. Sein Begehren wurde ihm hierauf in
so weit zugestanden/ daß er von einer Vesperglocke zur andern sicher
Gelait haben/ und er sich am S. Georgen Tage zu Grätz einfin-
den sollte/ wie auch geschah.

Nachdem er nun verhöret wurde/ und allerhand Ausflüchte/ sein
Thun zu rechtfertigen/ auf die Bahn brachte/ ruckte das Ende des
Termins herben/ welchen er von Kayser Friderichen zu verlän-
gern bate/ der sich mit seinen Rächten deswegen zu bereben vorsätz-
te/ und so lange damit verzogte/ bis die Zeit des Gelaites zu versieffen
begunnte/ dannhero Baumkircher vom Schloß herunter liefe/
sich auf sein Pferd schwangete/ und damit zum Thor hinaus zu rennen
trachtete.

D ij Weil

Information

Wiener-Moden-Zeitung und Zeitschrift für
Kunst schöne Literatur und Theater
(Jg.1.1816,1)

Schickh, Johann

Wien: Strauß; 1816-1849

Zitierlink:

<http://data.onb.ac.at/ABO/%2BZ136925106>

Weitere Bände/Exemplare

- [Jg.10.1825,1](#)
- [Jg.10.1825,2](#)
- [Jg.10.1825,3](#)
- [Jg.10.1825,4](#)
- [Jg.11.1826,1](#)
- [Mehr...](#)

Suche



Übersicht



Wiener-Moden-Zeitung

Jg. 10.

Nr. 100

Wien, 1825

Verlag von J. Strauß

am Hauptplatz

Nr. 100

Wien

Weitere Bände



Ansicht



Information

Des loblichen Fürstenthumbs Steyr Erbhuldigung, inn dem Fünfftzehenhundert vnnd zwaintzigsten, Auch nochuolgend des ainundzwainzigsten Jars bescheen, sambt handlung der Landtäg des selben mals gehalten, mit ettlichen ihren Freyhaitten, Landshanduest vnnd berüff der ordnung in gemainen handtierungen

Hofmann, Hans
Augsburg: Necker; 1523
Zitierlink:
<http://data.onb.ac.at/ABO/%2BZ182016604>

Suche

"Graz"

Übersicht



Navigation controls: 87 von 134, keine Treffer auf diesem Bild, Ansicht

XXXVIII

Wir als liebhaber der gerechtigkeit/gantz ab/ gebed vnd bestatent jnen auch Xate der gerechtigkeit die freyhait/das sy ire Sun vn töchter künfftiglich verheyraten sollen/wem sy wollen/Welcher Steyrer aber an Testament verschaidet/des selben Erbe oder nachstgefibten freünd / soll sein verlassen güt eruolgen . Ob sich begab / das zwischen Steyrern klag oder vnwillen entsprung / so soll der kampf oder beweysung durch kampf nit stat haben / sonnder es sollen geschickt zeügen sürgerstelt / vnnd nach der selben gezeugenuss die klag oder vnwillen entschiden werden. Aber die klagen über grund vnd boden/sollen von denen Richtern / so darzu gestelt / nach verhörung der zeügen sag/ mit rechtlicher ordnung nach der gerechtigkeit geendet werden. In Lehen gütern thun wir genglich ab die beschwärde/ so man gemaincklich neent anfall/ als das/so güten vnd Ersamen gewonhaitn widerwertig ist / Auf gnad vnser angebomen gütigkeit/ setzen wir in krafft diser freyhait/ das die töchter komen in die güter der vätter/wo sy nit haben mändlich Erben/dardurch der väter gedächtnus in denen sünen gemert wirdet. Weiter setzen wir / das die Lehen von ainander überkomen / wo sy gleich von dem herrn des Erdreichs für eigenthumb erkaufft wurden/ dem sollen gelassen werden/der die von recht des Lehen besitzt. Verrer vergünnen wir

Wer des abgestorbenen güt erbē sol.

Verpor des Kampfes.

Zürichten vmb grund vnd boden.

Anfall.

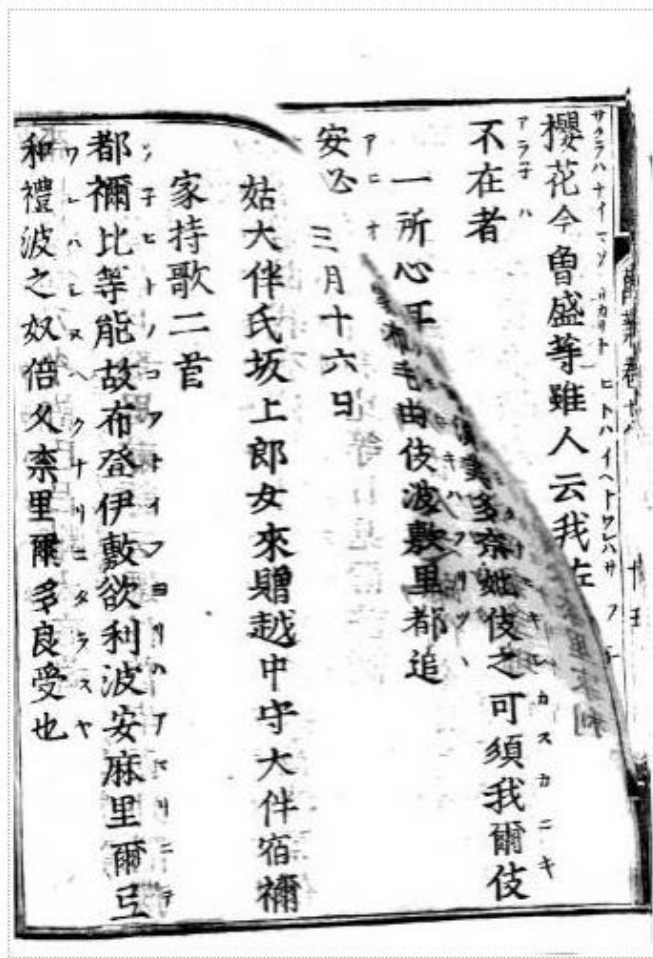
Es die töchter Erben mögen.

So der fürst Lehengüter kauft.

Information
Man'yōwakashū (Getr. Zählung)
[S.l. Kyōto]: Izumoji Izumi no Jō; 1709
Zitierlink:
<http://data.onb.ac.at/ABO/%2BZ172208207>

Suche

Übersicht



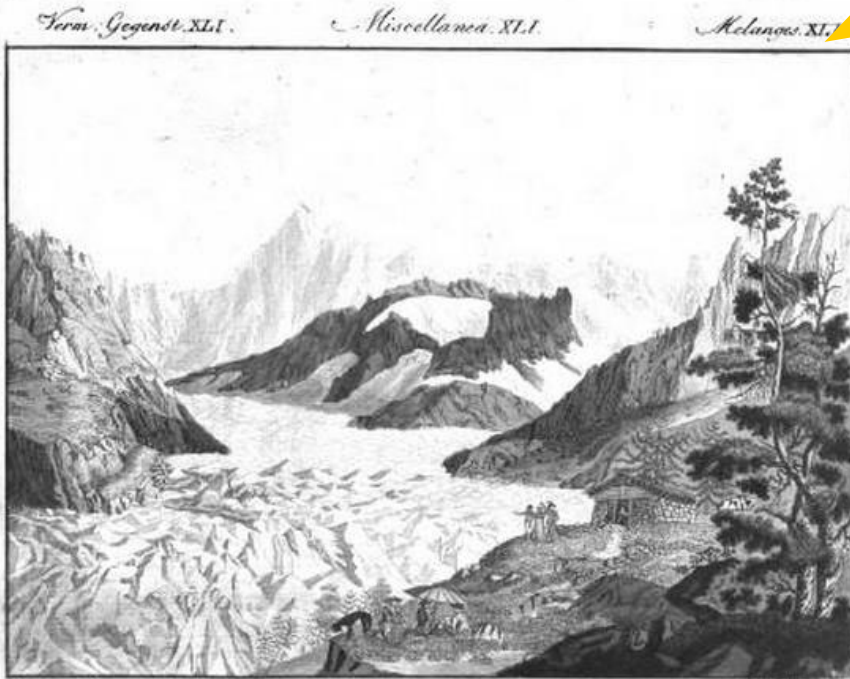


Anno CHRISTI *Saluatoris;*
1578.

[illegible]

HINWEIS: die Textdaten wurden mit einem automatisierten OCR-Verfahren ermittelt und weisen Fehler auf

... und doch, verschiedene Qualitäten



HINWEIS: die Textdaten wurden mit einem automatisierten OCR-Verfahren ermittelt und weisen Fehler auf

Bertuch, J F : Bilderbuch zum Nutzen und Vergnügen der Jugend; Wien 1805

Vermischte Gegenstände XLI.

T. VIII. No. 3.

ALPENGLÄTSCHER.

Das Eismeer am Montanvert.

Die Schweizeralpen sind die höchste Erhöhung der Erde in der alten Welt, und sie enthalten zugleich den höchsten Berg in den drej Welttheilen Europa, Asia und Africa, den *Montblanc*, der bloß von dem ungeheuern *Chimborasso* in Südamerika an Höhe übertroffen wird.

Eben diese hohen Schweizeralpen bieten uns wegen ihrer ausgezeichneten Lage die höchsten Naturschönheiten und die merkwürdigsten Gegenstände dar. Sie enthalten Berge, deren Gipfel über die Wolken steigen, und mit ewigem Schnee bedeckt sind; zwischen diesen Bergen befinden sich Seen, ungeheure Felsenabgründe, Thäler voll Eis, unzählige Quellen, welche in der Folge zu großen Flüssen werden; Bäche, die sich über Felsenwände herabstürzen, in der Luft verschwinden, und sich in Wasserstaub auflösen; fruchtbare Viehweiden, und andere Wunder der Natur. Einer der merkwürdigsten Gegenstände der Alpen sind ihre Glätscher.

Glätscher nennt man die ganz mit ungeheuren Eismassen angefüllten Alpenthäler, wel-

che theils höher zwischen den Felsenspitzen, und abhängig, theils tiefer und weiter am Fusse der höchsten Gebirge liegen. Einer der größten und merkwürdigsten Glätscher ist das sogenannte *Eismeer* am Berge *Montanvert*. Es ist ein langes ganz mit Eise ausgefülltes Thal, welches man von der Höhe des *Montanvert*, auf welcher hier die Hütte steht, übersehen kann, und einem Meere gleicht, dessen Fluthen plötzlich zusammen gefroren sind, aber nicht während des Sturms, sondern gleich nachher, da der Wind sich gelegt, die Wellen aber schon stumpf und abgerundet sind. Diese Eiswellen haben zwischen sich große und tiefe Querspalten, welche innerlich blau aussehen, und in welche ein Wanderer leicht hineinfallen kann.

Von beyden Seiten kommen zwischen den Bergspitzen noch kleinere Nebenglätscher herab, und im Hintergrunde erblickt man hier den großen ganz mit Schnee bedeckten *Jarassus*, welcher einer der höchsten Berge der Alpen ist. Der *Montanvert* selbst, auf welchem sich hier im Vorgrunde die Gesellschaft befindet, ist eine fruchtbare Alpentrieff, gleich neben diesen fürchterlichen Eismassen.

Vermischte Gegenstände Xbl. T. FUI. Ufo. S. ALPEJGLÄTSCHER.

Das Eismeer am Montanvert.

Die Schweizeralpen sind die höchste Erhöhung der Erde in der alten Welt, und sie enthalten zugleich den höchsten Berg in den drej Welttheilen Europa, Asia und Africa, den *Montblanc*, der bloß von dem Ungeheuern *Chimborasso* in Südamerika an Höhe übertroffen wird.

Eben diese hohen Schweizeralpen bieten uns wegen ihrer ausgezeichneten Lage die höchsten Naturschönheiten und die merkwürdigsten Gegenstände dar. Sie enthalten Berge, deren Gipfel über die Wolken steigen, und mit ewigem Schnee bedeckt sind; zwischen diesen Bergen befinden sich Seen, ungeheure Felsenabgründe, Thäler voll Eis, unzählige Quellen, welche in der Folge zu großen Flüssen werden; Bäche, die sich über Felsenwände herabstürzen, in der Luft verschwinden, und sich in Wasserstaub auflösen; fruchtbare Viehweiden, und andere Wunder der Natur. Einer der merkwürdigsten Gegenstände der Alpen sind ihre Glätscher.

Glätscher nennt man die ganz mit ungeheuren Eismassen angefüllten Alpenthäler, welche theils höher zwischen den Felsenspitzen und abhängig, theils tiefer und weiter am Fusse der höchsten Gebirge liegen. Einer der größten und merkwürdigsten Glätscher ist das sogenannte *Eismeer* am Berge *Montanvert**. Es ist ein langes ganz mit Eise ausgefülltes Thal, welches man von der Höhe des *Montanvert*, auf welcher hier die Hütte steht, übersehen kann, und einem Meere gleicht, dessen Fluthen plötzlich zusammen gefroren

Miscell. XLV.

T. VIII. No. 3.

MOLES GLACIALES ALPINAE.

Mare glaciale ad Montem viridem in Sabaudia.

Alpes Helvetiae et Sabaudiae summa habent orbis veteris cacumina, ex quibus maxime omnium prominet *Mons albus* in Sabaudia, quem solus ingens ille Chimborassus, qui in America meridionali attollitur, altitudine superat.

Harum alpium juga admirabili situ locorum, quascunque res aut visu amoenas aut memoratu dignas rerum natura continet, peregrinantibus spectandas proponunt. Hic eminent montes perpetuis nivibus tecti, quorum vertices supra nubes exsurgunt; in his autem ampli stagnant lacus, hiant immanes rupium voragine, desidunt valles glacie ad summum oppletæ, promanant undique innumeri fontes, qui exinde collecti validos in amnes intumescunt; torrentes, de saxeis parietibus ex alto dejecti, cadendo per aërem dissipantur, tenuissima adspersione aqueos in pulvisculos evanescentes; alibi **pinguis** luxuriantur pascua; varia ubique naturæ miracula spectatorem invitant.

Inprimis memoratu dignae sunt moles illae glaciales, quos alpini *Glaciaris* vocant,

His molibus implentur quaedam valles alpinae, partim superius inter scopulos declivi latitudine expansae, partim inferius ad radices montium altissimorum projectae. Plurimum inter has glaciei congeries ob magnitudinem formamque mirabilem fama celebratur *Mare glaciale*, *Monti viridi* conterminum. Sic profundum gominant vallem, quam praesens tabula refert, oppletam glacie. Ex *Monte viridi*, cui nostra in iconetugurium superimpositum est, prospectus patet in hanc vallem, haud absimilem mari, cujus fluctus, non cum saeviret procella, sed ubi desaevit ventus, languentibus undis ac jamjam rotundioribus dicas subito frigore congelavisse. Hos inter fluctus congelatos ex obliquo patescunt rimae grandes ac profundae, intus caeruleae, incautius accedentem devoraturae.

Ab utroque vallis latere inter montium juga prominent aliae, sed minores, glaciei congeries; interiori autem in regione conspicitur ingens *Jurassus*, nivibus tectus, altissima alpium montibus adaumerandus. Ipse *Mons viridis*, cujus in latere anteriori societas, quam vides, peregrinantiam commoratur, quamvis horridae huic glaciei vicinus, pascuis tamen laetissimis gaudet.

Miscell. JĭLI..

T: FUT. III. 3ĭ.

MOLES GLACIALES ALPINAE.

Mare glaciale ad Montem viridem in Sabaudia.

Alpes Helvetiae et Sabaudiae summa habent erbis veteris cacumina, ex quibus maxime omnium prominet Mōns albus in Sabaudia, quem solus ingens ille Chimborassus, qui in America meridionali attollitur, altitudine supera t..

Harum alpium juga admirabili situ Iocorum, quascunque res aut visu amoenas aut memoratu dignas rerum natura contmet, peregrinantibus spectandas proponunt. Hic eminent montes perpetuis nivibus tecti, quorum vertices supra nubes exsurgunt; in his autem ampli stagnant lacus, hiant immanes rupium voragines, desidunt valles glacie ad summum oppletæ, promanant undique innumeri fontes, qui exinde collecti validos in amnes intumescunt }. torrentes, de saxeis parietibus ex alto dejecti, cadendo per aerem dissipantur, tenuissima adspersione aqueos in pulvisculos evanescentes; alibi pinguis luxuriantur pascua; varia ubique naturæ miracula spectatorem invitant.

Inprimis memoratu dignae sunt moles illae glaciales, quos alpini Glaciaris vocant. His molibus implentur quaedam valles alpinae, partim superius inter scopulos declivi latitudine expansae, partim inferius ad radices montium altissimorum projectae.. Plurimum inter has glaciei congeries ob magnitudinem formamque mirabilem fama celebratur Mare glaciale, Monti viridi conterminum.. Sic profundum n,ominant vallem, quam praesens

*Elegystárjak XLI.**VIII. Dar. 3. Szám.*

AZ ÁLPESI JEGESEK.

A' Jégtenger Montavernél.

A' Helvétziai havasok a' régi világnak legmagasabb domborodásai, ezek között van a' föld három részeinek Európának, Ázsiának, 's Afrikának legmagasabb hegye az úgy nevezett *Mont-blanc* is, melyet egyedül csak déli Amérikába a' reitenetes *Tsimboráfzfo* halad meg.

Ugyan ezen Havasok, kiválogatott különös fekvéseikre nézve a' természet legtsudálatra méltóbb szépségeivel 's legnevezetesebb tárgyaival gyönyörködtetik az érzékenységeket. Vannak ezeken oly hegyek, melyeknek tetejek a' felhőkön fellyül emelkedett, ezek között a' hegyek között helyel helyel tavak, feneketlen köfzikla barlangok, jég völgyek, számlálhatatlan források találhatnak, olyan források melyek tovább nagy folyókat formálnak; patakok, melyek a' tornyos köfzálókról lebukván, a' levegőbe eltűnnek és vízporrá válnak; kővér legelők és más tsuda dolgai a' természetnek. Az Alpesi legnevezetesebb tárgyak közül valók a' *Jegesek*.

Jegeseknek nevezetnek azok a' teménten jég darabokkal megtöltött völgyek, melyek részént a' köfzálóknak közeit 's hajlatait, részént pedig a' hegynek allját foglalják el. Egy a' legmegjegyzésre méltóbb jegesek közül az úgy nevezett *Jégtenger Montavernél*. Ez egy hófzfú jéggel megtöltött völgy, melyet Montavernék tetejéről, a' hol egy gunyhó is van, egészen beláthatni. Hasonló ez ahoz a' tengerhez, mely nem éppen a' fzelvfek háborgása alatt, hanem akkor állott vagy fagyott öfzve, mikor a' fzel kezdván tsendesedni annak vize féldomború hullámokba tolakodott. Ezen jég hullámok között helyel helyel belőlről kékelő rézsut nyílások vannak, melyekbe a' vizgálódó könnyen elbukhatik.

Két oldalról kisebb jegesek ereszkednek le, az általlenbe lévő oldalon pedig az örök hóval borított *Jarassus* egyik Alpesi legmagasabb hegyet szemléltetni. A' Montaver pedig maga, a' melynek elején a' szemléldő társaság láttatik, kies és termékeny Alpesi legelő, mindjárt ezen ifzonyú jég vidék izomfűdségába.

Elegy estirgyak Xbl. * VIII. Dar, 3. Szám.

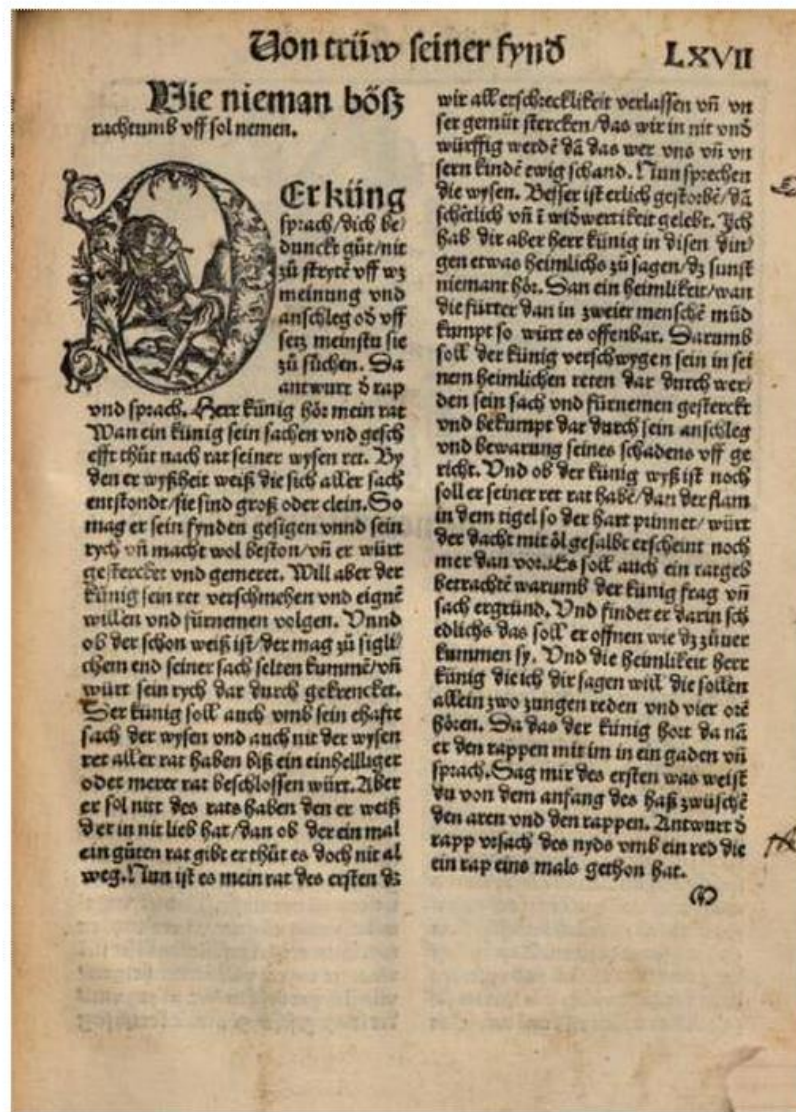
A Z ALP ESI JEGESEK.

A' Jégtenger Montavernél.

.A*Helvétziaihavasok a'regi világnaklegmagasabb domborodásai r ezek között van a' föld három részeinek Európának , Ázsiának, 's Afrikának legmagasabb hegye az úgy nevezett Mont-blanc is, melyet egyedül tsak déli Amérikába a' rcitenetes Tsimboráfzfo halad meg,

Ugyan ezen Havasok, kiválogatolt különös fekvéseikre nézve a' természet legtsudálatra méltóbb ízéségeivel 's legnevezetesebb tárgyaival gyönyörködtetik az érzékenységeket. Vanoak ezéken oily hegyek, melyeknek tetejek a'felhőkön fellyül emelkedett, ezek között a' hegyek között helyel helyel tavak, feneketlen köfzikla barlangok, jég völgyek , számlálhatatlan források találhatnak, olyan források me'.Jyek tovább nagy i'olyókat formálnak; patakok, melyek a' tornaos koTzálokiól lebukván, a' levegőbe eltűnnek és vízporrá válnak; kővér legelők és más tsudadolgai a' természetnek. AzÁlpesi legnevezetescbb tárgyak közül valók 3?Jcgusek. Jegesehneh nevezetnek azok a' teménten jég darabokkal megtöltött völgyek , melyek részént a' köfzálóknak közeit 's hajlatait, rérzént pedig a' hegynek allját foglalják el. Egy a' legmegjegyzésre méltóbb jegesek közül az úgj' nevezett Jégtenger Montavernél. Ez egy hófzfú jéggel megtöltött völgy , melyet Montavernék tetejéről,, a' hol egy gunyhó is van, egészen beláthatni. Hasonló ez ahoz a* tengerhez, mely nem éppen a' fzelvfek háborgása alatt, hanem akkor állott vagy fagyott öfzve, mikor a' fzel kezdván tsendesed-

**Österreichische
Nationalbibliothek**



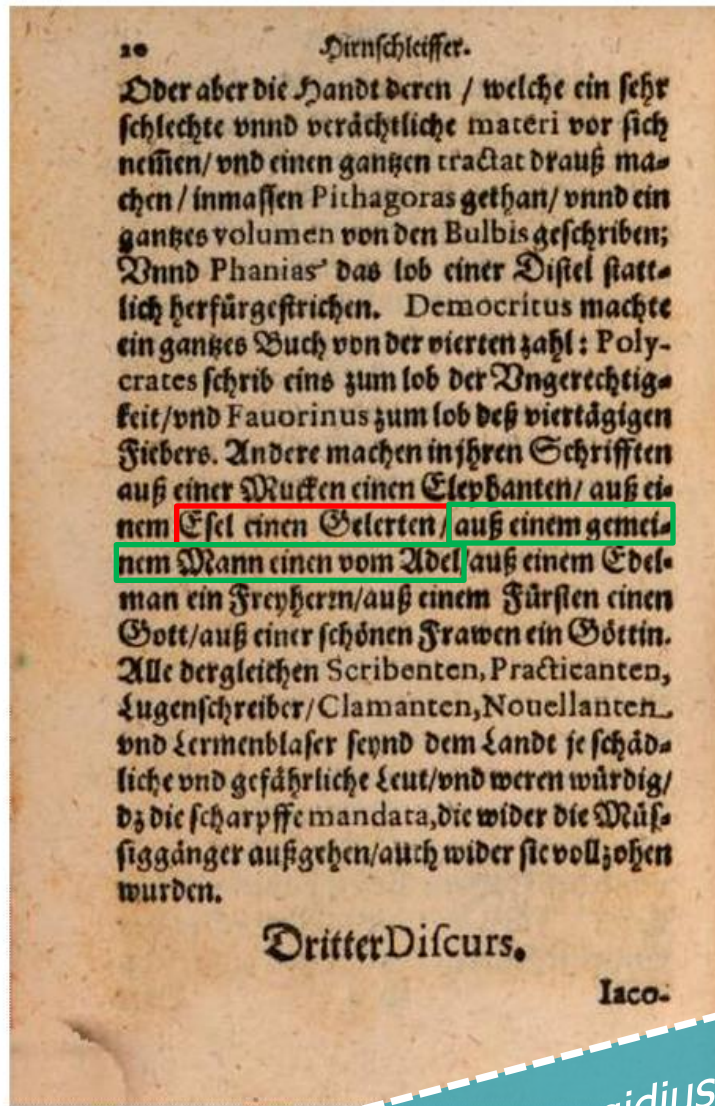
U011 criiiv "einer _nme
Wie meinen böllz
racstnmsvtff0lnemen. .

"F : * fpiacs/?ici sei
* " f* *. dnncr'r güt/nit
l - ' : zii fig-te* vffwz
x . ' i i f meinung vnd
- . anfoßleg 05 vff
- * " feiz meinfkn fie
"- zi'i fließen. Ia
- ' ' antwrt d rap

vnd fpiacs. Zerr Einig sie mein rat
"Man ein ?iinig fein fachen vnd gefcß
efft cöiit naci? rat feiner wyfen rer. By
?en er wyßöeit weiß ?ie fiel? after fat?
entfk0ndt/fie ['ind groß oder clein.Sd
mag er fein fynden gefigen ?nnd reed' sn macß't wol
sefldn/vn er wurc
gez-ferner vnd gemeret. win aser ?er
Einig [ein rec verfcsmessen vnd eigne
wiilen vnd [ijrnemen folgen. ?Zi-ind
08 ?er fciion weiß (fi)?er mag zu [iglxj
ciJem end feiner fact? felten Summe/vn
wiirt fein ryys ?ar ?uns geErencE'et.
'Ser ?iinig foik" ancß vini* fein esafte
"Ms '3er wrfen vnd anti? nitOer wyfen
rec aik'er rat saßen siß ein einselliger
oder merer rat b'efciiloffen würt.2(Cek
er f 0l nitt ?es race saßen c.len er weiß
?er in nit lies Cat/Ban 08 c6er ein mn(
ein güten ratgist er tsi'it es ?des nit al-
wegWiin iii es mein rat ?es erfken ?z

xxx/II

Beispiel Fraktur (schlechte Qualität): *Dis ist das buch der wysheit
der alten wysen von geschlecht der welt.*; Bidpai, Person der Antike
oder des Mittelalters; Straßburg: Grüninger; 1501 Hainrich; 1618



1. Hirnfekyleifier.

Oder arm-die .hunde deren / welche ein [ehr [eh/lechte vnnnd verächtliche maceri vor sich neffien/ vnd einen gangen tractat drauß ma- chen / inma (sen Pithagoras gefhan/ vnnnd ein V anBee volumcn von den Bulbis esczriben; nd Ph ania-s" das lob cintr isel [latt- lidz herfür gestrichen. Dcmocricus machtc ein ganzes Buch von der vierten zahl : Poly- crates [ehr-ib eine zum lob dei-Vngerechtig- feic/vnd Fauorinus zum [ob des viertelgigen Fieber-s. "Andere machen inihren S czriffcn ans einer Muck-en einen Elephane auß ei- nem Esel einen Gummi auß einem gemei- nem Mann einen vom Adel/ auß einem Eva- man ein Freyheren/ auß einem Für-urn einen Gott/ann einer [ebenen stenten ein Göttin. Alle dergleichen ScribentcmPracticantcu, clpgtnffhrtibcr/ClamantincchNoucllantctL. vnd Zermenblaser feynd dem .Lande je schäd- lithe vnd gefährliche Heut/nnd weren wüldig/ dz die scharpffe mandacpie wider die Muz. figgängernußgehen/ auch wider sie vollzohen wurden.

Orittcr Discurs.
lace."

HINWEIS: die Textdaten wurden mit einem automatisierten OCR-Verfahren ermittelt und weisen Fehler auf

:Albertinus, Aegidius Hiren-schleifer. (Mit Kupfern); München: Niclas Hainrich; 1618

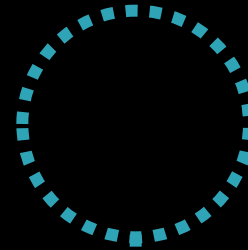
Austrian Books Online

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

Austrian Books Online

Working with the Data

Buchlogistik> Digitization



Daten-Download

Storage

QA

Access

ADOCO
(Austrian Books Online
Download & Control)

Workflow in ADOCO



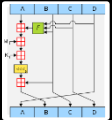
Download package via HTTP



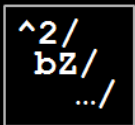
Decrypt with gnupg



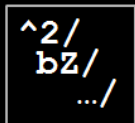
Unzip tarball



Md5 sum



Store to pairtree



Unified Access → Pairtree (Symlinks)



Update metadata

Volume

Average per Volume (~Book): 101 MB

$$101 \text{ MB} * 600.000 = 60 \text{ TB}$$



Image courtesy of The University of Pennsylvania and Michel T. Huber. www.fi.edu

big data ...

Datenspeicherung & Access

- Datenspeicherung: inhouse
- Daten redundant gespeichert
- Access-Kopien on-the-fly generiert

Download und Speicherung

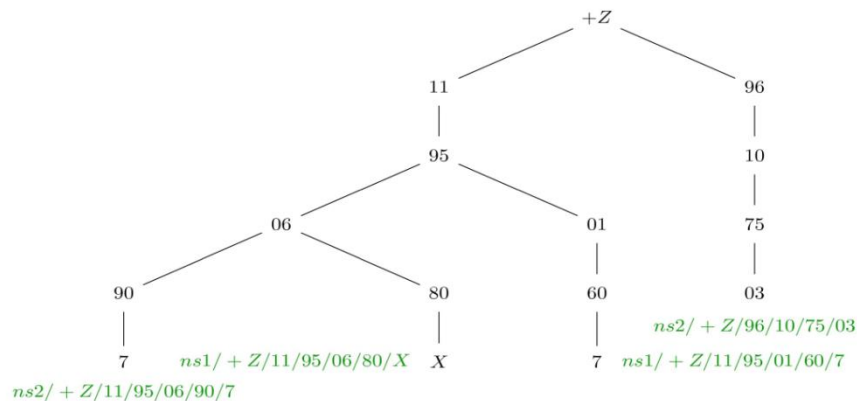


ADOCO

JPEG2000
HOCR
METS
TXT



ABO NAS-Speicher
Pair Tree-Algorithmus
ca. 60 TB



Pair Tree: ABO NAS

+Z156941203

^2/
bz/
15/
69/
41/
20/
3/abo/

ONB_+Z156941203.xml

00000001.html

00000001.jp2

00000001.txt

→ <https://confluence.ucop.edu/display/Curation/PairTree>

Datenorganisation

→ METS (Metadata Encoding & Transmission Standard)

<http://loc.gov/standards/mets/>

→ MARC/XML / MODS

→ PREMIS

→ GBS specific metadata

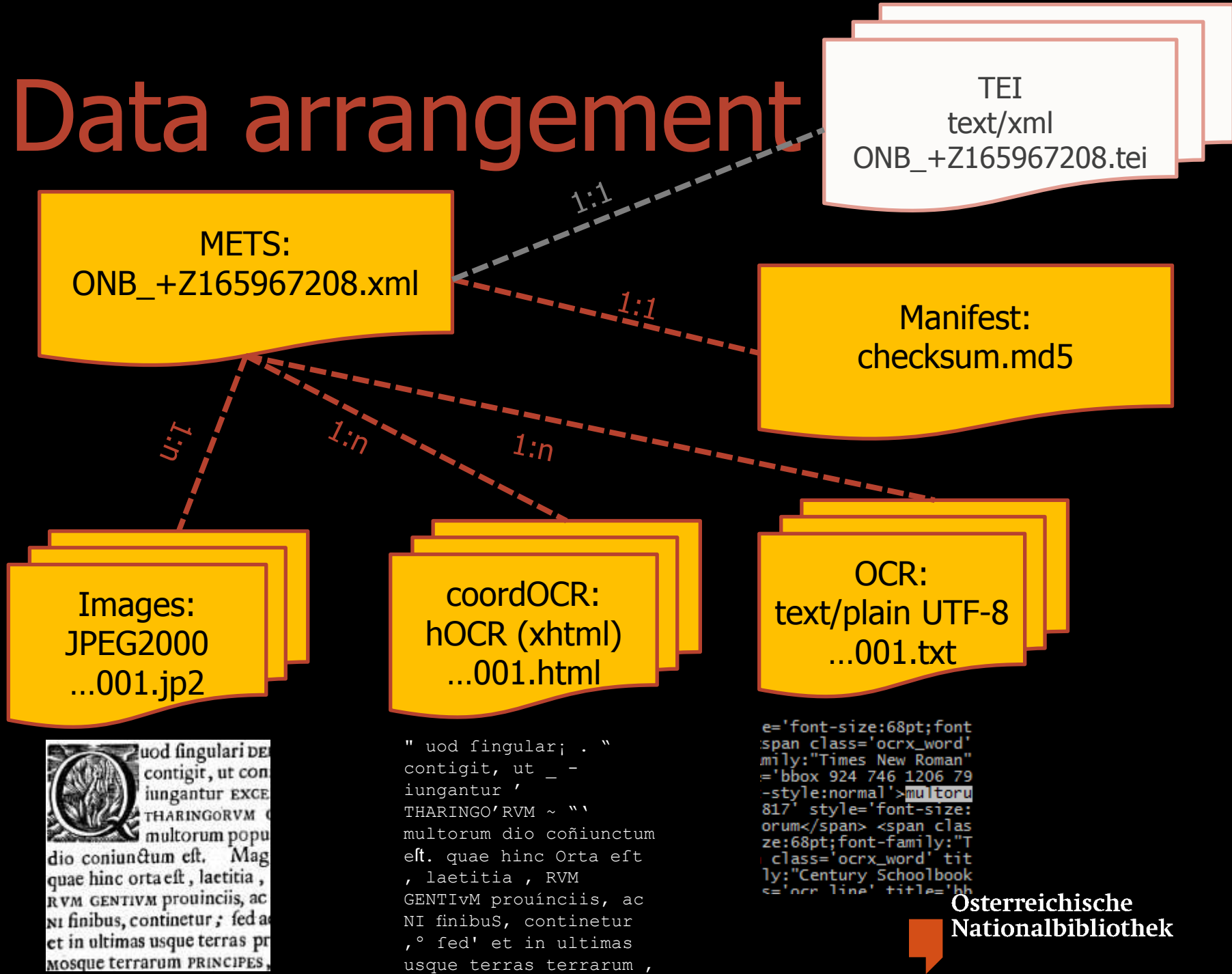
→ Images (JPEG2000)

→ OCR Daten

→ Coordinated OCR

→ plain TXT

Data arrangement



METS

→ Reference:

<http://www.loc.gov/standards/mets/METSOverview.v2.html>

→ Namespaces:

xmlns:METS="http://www.loc.gov/METS/"

xmlns:xlink="http://www.w3.org/1999/xlink"

xmlns:gs="http://books.google.com/gbs"

xmlns:PREMIS="info:lc/xmlns/premis-v2"

xmlns:marc="http://www.loc.gov/MARC21/slim"

METS

→ Structure

→ METS:mets

- METS:metsHdr
- METS:dmdSec
- METS:amdSec
- METS:fileSec
- METS:structMap

METS:metsHdr

```
<METS:metsHdr ID="HDR1" CREATEDATE="2013-08-08T02:56:26" RECORDSTATUS="NEW">  
  <METS:agent ROLE="CREATOR" TYPE="ORGANIZATION">  
    <METS:name>Google</METS:name>  
  </METS:agent>  
</METS:metsHdr>
```

METS:dmdSec

```
<METS:dmdSec ID="DMDMARC" xmlns="http://www.loc.gov/MARC21/slim">
<METS:mdWrap MIMETYPE="text/xml" MDTYPE="MARC" LABEL="Source MARC record">
<METS:xmlData>
<record xmlns="http://www.loc.gov/MARC21/slim"^M
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"^M
xsi:schemaLocation="http://www.loc.gov/MARC21/slim^M
http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd">^M
  <leader>      nam  2200000 u 4500</leader>^M
  <controlfield tag="001">LZ01632926</controlfield>^M
  <controlfield tag="008">                                |||   ger</controlfield>^M
  <datafield tag="041" ind1="0" ind2="7">^M
    <subfield code="a">ger</subfield>^M
    <subfield code="2">ISO 639-2</subfield>^M
  </datafield>^M
  <controlfield tag="007">zz|||||||||||||||||||</controlfield>^M
  <datafield tag="100" ind1="1" ind2=" " >^M
    <subfield code="a">Schröckh, Johann Matthias</subfield>^M
  </datafield>^M
  <datafield tag="700" ind1="1" ind2=" " >^M
    <subfield code="a">Tzschirner, Heinrich Gottlieb</subfield>^M
  </datafield>^M
  <datafield tag="245" ind1="0" ind2="0">^M
```

METS:amdSec

```
<METS:amdSec>
  <METS:techMD ID="TMD1">
    <METS:mdWrap LABEL="production notes" MDTYPE="OTHER" OTHERMDTYPE="Google Production Notes">
      <METS:xmlData>
        <gbs:productionNotes>
          <gbs:badPages>75.59</gbs:badPages>
          <gbs:missingPages>0.26</gbs:missingPages>
          <gbs:tightBoundPages>1.46</gbs:tightBoundPages>
        </gbs:productionNotes>
      </METS:xmlData>
    </METS:mdWrap>
  </METS:techMD>

  <METS:techMD ID="IMAGE_METHOD0">
    <METS:mdWrap LABEL="method of image production" MDTYPE="OTHER" OTHERMDTYPE="Scan Details">
      <METS:xmlData>
        <gbs:imageMethod scanner="PHOTOSTATION_V1">non-destructive</gbs:imageMethod>
      </METS:xmlData>
    </METS:mdWrap>
  </METS:techMD>

  <METS:techMD ID="CALIB_8801478227875552005">
    <METS:mdWrap LABEL="calibration target" MDTYPE="OTHER" OTHERMDTYPE="Calibration target">
      <METS:xmlData>
        <gbs:focusParameters threshold="40%">
        </gbs:focusParameters>
      </METS:xmlData>
    </METS:mdWrap>
  </METS:techMD>
```

METS:fileSec

```
<METS:fileSec>
  <METS:fileGrp ID="FG1" USE="image">
    <METS:file ID="IMG00000001" MIMETYPE="image/jp2" SEQ="00000001" CREATED="2011-04-08T00:00:00" SIZE="167133" CHECKSUM="06abca78fadf783f2d44e59352b03e49" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-1" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000001.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
    <METS:file ID="IMG00000002" MIMETYPE="image/jp2" SEQ="00000002" CREATED="2011-04-08T00:00:00" SIZE="95542" CHECKSUM="1c7b157b5af73ea33c63c4567423337f" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-2" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000002.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
    <METS:file ID="IMG00000003" MIMETYPE="image/jp2" SEQ="00000003" CREATED="2011-04-08T00:00:00" SIZE="70736" CHECKSUM="ef3614a1f99a72ab31251f0f212db127" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-3" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000003.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
    <METS:file ID="IMG00000004" MIMETYPE="image/jp2" SEQ="00000004" CREATED="2011-04-08T00:00:00" SIZE="25770" CHECKSUM="2b6df659fd007b6576cc16979bdb0e8d" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-4" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000004.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
    <METS:file ID="IMG00000005" MIMETYPE="image/jp2" SEQ="00000005" CREATED="2011-04-08T00:00:00" SIZE="20889" CHECKSUM="5eeb401f49b74d27e08b4f310f7afbd5" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-5" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000005.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
    <METS:file ID="IMG00000006" MIMETYPE="image/jp2" SEQ="00000006" CREATED="2011-04-08T00:00:00" SIZE="488810" CHECKSUM="21b88aafe9d7145583a1e2a606a288d4" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-6" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000006.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
    <METS:file ID="IMG00000007" MIMETYPE="image/jp2" SEQ="00000007" CREATED="2011-04-08T00:00:00" SIZE="547690" CHECKSUM="7141d6473d18a1f3e40ad9e001ce5703" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-7" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000007.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
    <METS:file ID="IMG00000008" MIMETYPE="image/jp2" SEQ="00000008" CREATED="2011-04-08T00:00:00" SIZE="8304" CHECKSUM="37f2f8f9aee96f09c782da162a50d47e" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-8" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000008.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
    <METS:file ID="IMG00000009" MIMETYPE="image/jp2" SEQ="00000009" CREATED="2011-04-08T00:00:00" SIZE="36729" CHECKSUM="cf46ca39d3b7938171e596b6a9e508f0" CHECKSUMTYPE="MD5" OWNERID="31525197396353934-9" ADMID="IMAGE_METHOD0">
      <METS:FLocat xlink:href="00000009.jp2" LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM"/>
    </METS:file>
  </METS:fileGrp>
</METS:fileSec>
```

METS:structMap

```
<METS:structMap ID="SM1" TYPE="physical">
  <METS:div TYPE="volume" ADMID="TMD1 PD1 S1">
    <METS:div TYPE="page" ORDER="1" ADMID="FRONT_COVER IMPLICIT_PAGE_NUMBER">
      <METS:fptr FILEID="IMG000000001"/>
      <METS:fptr FILEID="TXT000000001"/>
      <METS:fptr FILEID="HTML000000001"/>
    </METS:div>
    <METS:div TYPE="page" ORDER="2" ADMID="IMPLICIT_PAGE_NUMBER">
      <METS:fptr FILEID="IMG000000002"/>
      <METS:fptr FILEID="TXT000000002"/>
      <METS:fptr FILEID="HTML000000002"/>
    </METS:div>
    <METS:div TYPE="page" ORDER="3" ADMID="IMPLICIT_PAGE_NUMBER">
      <METS:fptr FILEID="IMG000000003"/>
      <METS:fptr FILEID="TXT000000003"/>
      <METS:fptr FILEID="HTML000000003"/>
    </METS:div>
    <METS:div TYPE="page" ORDER="4" ADMID="IMPLICIT_PAGE_NUMBER">
      <METS:fptr FILEID="IMG000000004"/>
      <METS:fptr FILEID="TXT000000004"/>
      <METS:fptr FILEID="HTML000000004"/>
    </METS:div>
    <METS:div TYPE="page" ORDER="5" ADMID="IMPLICIT_PAGE_NUMBER">
      <METS:fptr FILEID="IMG000000005"/>
      <METS:fptr FILEID="TXT000000005"/>
      <METS:fptr FILEID="HTML000000005"/>
    </METS:div>
  </METS:div>
</METS:structMap>
```

METS:amdSec

→ METS:techMD

- „production notes“ (badPages, missing Pages, tightBoundPages)
- „method of image production“
- „calibration target“
- Definition of gbs:pageTag

METS:amdSec

→ METS:digiprovMD

- „production notes“ (badPages, missing Pages, tightBoundPages)
- „method of image production“
- „calibration target“
- Definition of gbs:pageTag

METS:amdSec

- METS:sourceMD

- Source library information

- METS:digiprovMD

- PREMIS:premis

- representation

- scanning date

- processing date

- analyzed date

- rubbish

hOCR



https://docs.google.com/document/d/1QQnIQtvAC_8n92-LhwPcjtAUFwBlzE8EWnKAxlgVf0/

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>OCR Output</title>
    <meta http-equiv="content-type" content="text/html; charset=utf-8" />
    <meta http-equiv="content-style-type" content="text/css" />
    <meta name="ocr-capabilities" content="ocr_page ocr_par ocr_block ocrx_word ocrp_lang ocr_line" />
    <meta name="ocr-system" content="ABBY fre-9.0.0.6953" />
    <meta name="ocr-number-of-pages" content="1" />
  </head>
  <body bgcolor="#ffffff">
    <div class="ocr_page" title="bbox 0 0 1141 2068;x_wconf 64;ppageno 111;ocrp_lang unknown">
      <div class="ocrx_block" title="bbox 16 217 288 265">
        <p class="ocr_par" title="bbox 16 217 288 265;x_wconf 46" style="font-size:10pt;font-family:"Times New Roman";font-style:normal">
          <span class="ocr_line" title="bbox 16 217 288 265;x_wconf 28">
            <span class="ocrx_word" title="bbox 16 226 85 265">Mit</span>
            <span class="ocrx_word" title="bbox 98 222 212 261">worden;</span>
            <span class="ocrx_word" title="bbox 242 217 288 256">her</span>
          </span>
        </p>
      </div>
      <div class="ocrx_block" title="bbox 16 264 291 360;x_wconf 46">
        <p class="ocr_par" title="bbox 16 264 291 360;x_wconf 46" style="font-size:10pt;font-family:"Times New Roman";font-style:normal">
          <span class="ocr_line" title="bbox 16 264 291 316;x_wconf 50">
            <span class="ocrx_word" title="bbox 16 276 70 316">Wft.</span>
            <span class="ocrx_word" title="bbox 109 264 291 310">Zwar""</span>
          <br />
        </span>
        <span class="ocr_line" title="bbox 16 309 198 360;x_wconf 37">
          <span class="ocrx_word" title="bbox 16 322 53 360">!in,</span>
          <span class="ocrx_word" title="bbox 68 315 113 357">diß</span>
          <span class="ocrx_word" title="bbox 120 313 160 346">bie</span>
          <span class="ocrx_word" title="bbox 171 309 291 350">P</span>
        </span>
      </p>
    </div>
```

Using the data, locally

```
scape@scape-desktop:~/Hackathon/abo$ pwd
/home/scape/Hackathon/abo
scape@scape-desktop:~/Hackathon/abo$ find . | head -n 14
.
./barcode_ids
./barcode_ids/barcodeids.txt
./metadata.csv
./pairtree
./pairtree/^2
./pairtree/^2/b2
./pairtree/^2/b2/18
./pairtree/^2/b2/18/06
./pairtree/^2/b2/18/06/66
./pairtree/^2/b2/18/06/66/80
./pairtree/^2/b2/18/06/66/80/2
./pairtree/^2/b2/18/06/66/80/2/abo
./pairtree/^2/b2/18/06/66/80/2/abo/00000648.jp2
scape@scape-desktop:~/Hackathon/abo$
```

→ <https://www.dropbox.com/s/zpb7jzti0f8gsxn/pairtree.sh>

Using the data, cluster

Paths:

/user/onbfue/input/abo/paths/mets/abo_mets_file_paths.txt

/user/onbfue/input/abo/paths/text/abo_text_file_paths.txt

/user/onbfue/input/abo/paths/html/abo_html_file_paths.txt

Data:

/user/onbfue/input/abo/data/html/seqfiles (page level)

/user/onbfue/input/abo/data/text/seqfiles (book level)