

Digital Transformation of Cultural Heritage: Exploring and Curating Historical Resources

**Large surveys with small telescopes: Past, Present, and Future
(Astroplate III)**



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Bamberg
11.-13.03.2019

Prof. Dr. Andrea Rapp
TU Darmstadt
@RappAndrea
@DHDarmstadt
rapp@linglit.tu-darmstadt.de

Check-in



TECHNISCHE
UNIVERSITÄT
DARMSTADT

The Turing (@turinginst) twitterte um 4:17 nachm. on Di., März 05, 2019:

“All data is historical data: the product of a time, place, political, economic, technical, & social climate. If you are not considering why your data exists, and other data sets don’t, you are doing data science wrong”. - @melissaterras
#TuringLectures <https://t.co/mjjGwl7GZu>
<https://t.co/zKrQYIYAk0>

<https://twitter.com/turinginst/status/1102951379507789826?s=09>

The Turing
@turinginst

“All data is historical data: the product of a time, place, political, economic, technical, & social climate. If you are not considering why your data exists, and other data sets don’t, you are doing data science wrong”. -
@melissaterras
#TuringLectures
youtube.com/watch?v=bf6DsJ...
Translate Tweet



4:17pm · 5 Mar 2019 · Twitter Web Client

Table of Contents



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Check in
- The history of Digital Humanities: a brief look
- Humanities research cultures & Cultural Heritage
- Humanities research data
- Example I: Organizing Variations of Hieratic Script
- Example II: Letters of Brothers Grimm, Savigny & Bang
- Conclusion

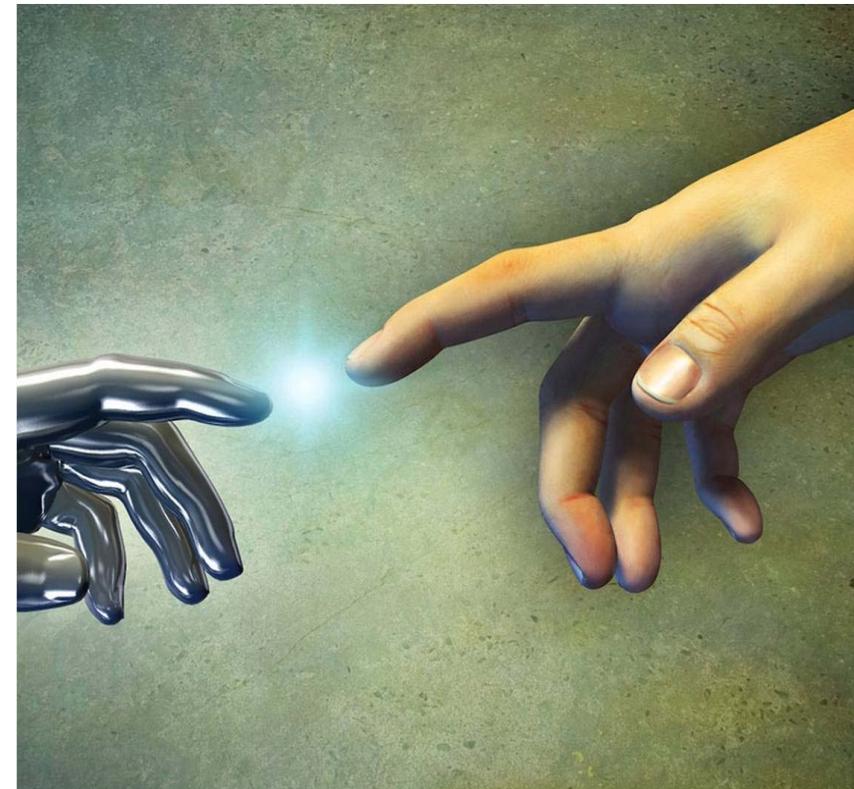
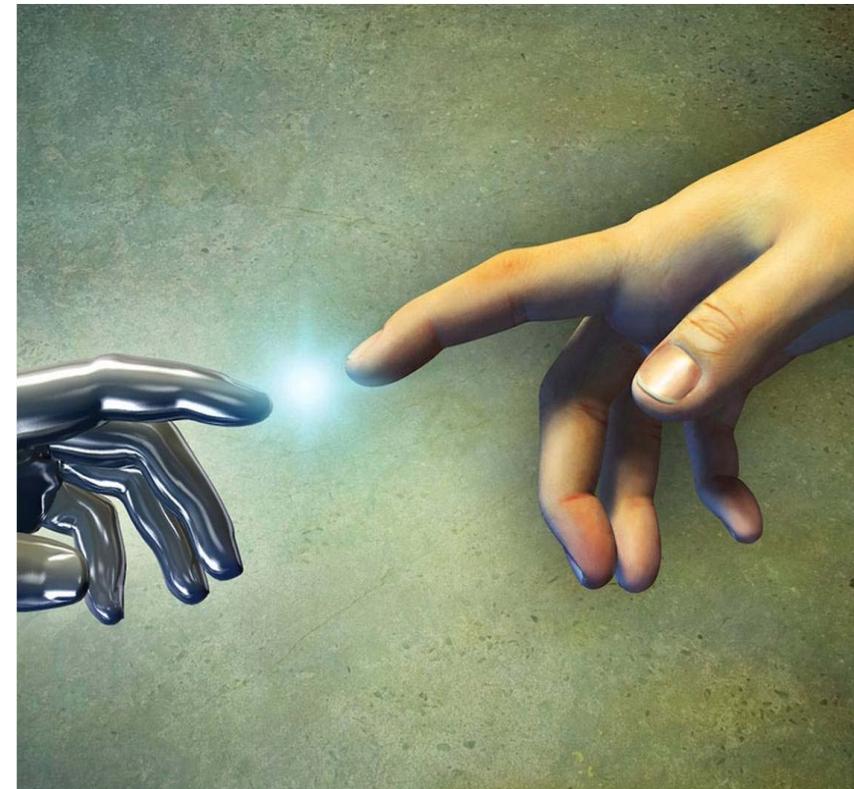


Table of Contents



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Check in
- **The history of Digital Humanities: a brief look**
- Humanities research cultures & Cultural Heritage
- Humanities research data
- Example I: Organizing Variations of Hieratic Script
- Example II: Letters of Brothers Grimm, Savigny & Bang
- Conclusion

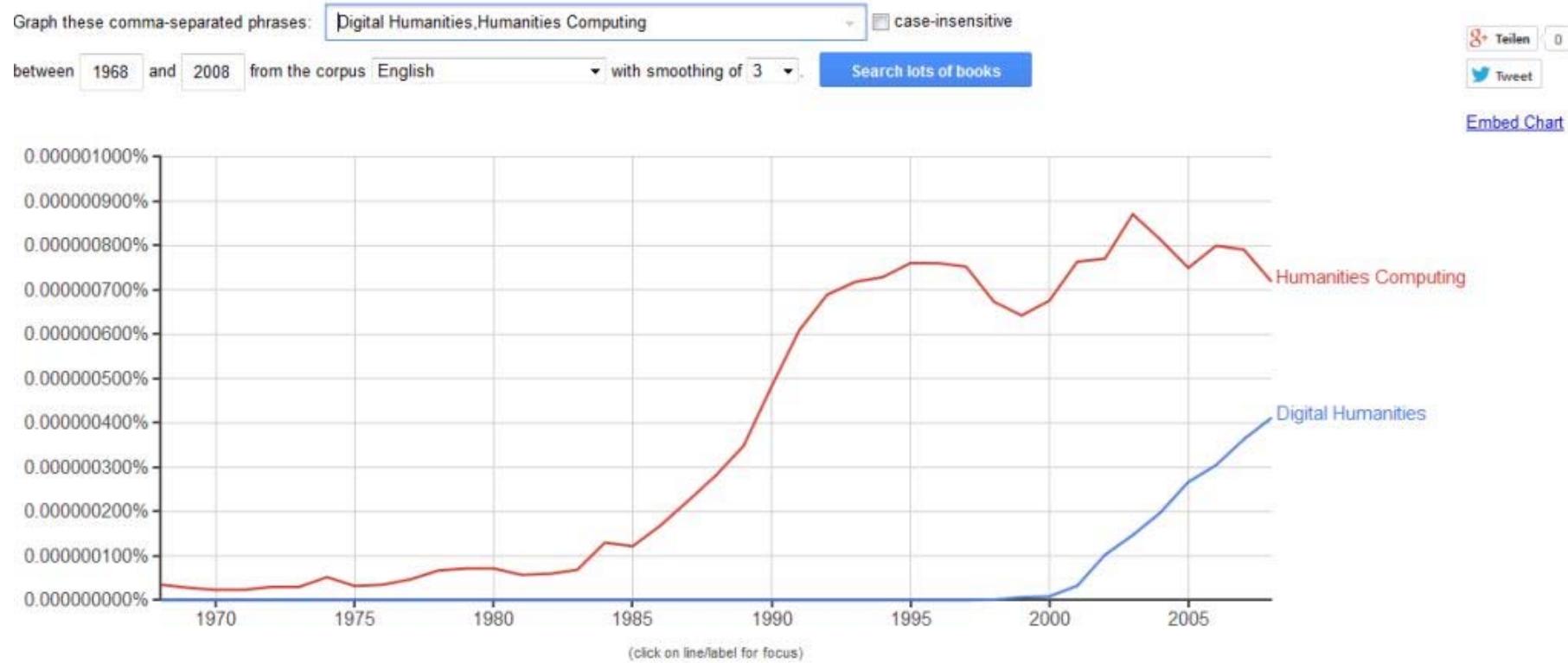


The History of Digital Humanities: Humanities Computing



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Google books Ngram Viewer



The History of DH: Humanities Computing



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Humanities computing is precisely the automation of every possible analysis of human expression (therefore, it is exquisitely a "humanistic" activity), in the widest sense of the word, from music to the theater, from design and painting to phonetics, but whose nucleus remains the discourse of written texts.

During World War II, between 1941 and 1946, I began to look for machines for the automation of the linguistic analysis of written texts. I found them, in 1949, at IBM in New York City. Today, as an aged patriarch (born in 1913) I am full of amazement at the developments since then; they are enormously greater and better than I could then imagine. *Digitus Dei est hic!* The finger of God is here!



Father Roberto Busa SJ: preface in: *A Companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens, John Unsworth. Oxford: Blackwell, 2004.
<http://www.digitalhumanities.org/companion/>

The History of DH: Index Thomisticus



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Starting with punchcards: 13 mio cards, 1 for each word with 12 lines of context on the backside: 90 m long, 1.20 m height, 1 m depth, weight 500 tons
- Magnetic Tapes: 1.800 tapes, each 2.400 feet long, all in all 1.500 km, way Paris-Lissabon or Milano-Palermo
- Publication: 20 tapes as setting copy, 20 mio lines, 65.000 pages, 56 items in encyclopaedec format = the printed *Index Thomisticus*
- Digital publication began 1987 with preparing data for a CD-ROM edition published in 1992, 1.36 GB
- Today: www.corpusthomisticum.org

Roberto Busa with Pope Paul VI presenting the printed edition



The History of DH: The Busa Archives



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Busa's female operators
(here Livia Canestraro)



- Melissa Terras: <http://melissaterras.blogspot.de/2013/10/for-ada-lovelace-day-father-busas.html>
- Julianne Nyhan: <http://archelogos.hypotheses.org/127>

The History of DH: www.corpusthomisticum.org



TECHNISCHE
UNIVERSITÄT
DARMSTADT

CORPUS THOMISTICUM
INDEX THOMISTICUS
by Roberto Busa SJ and associates
web edition by Eduardo Bernot and Enrique Alarcón
English version

Search:

[concordances](#) [terms](#) [works](#) [options](#) [new search](#)

Instructions:

You may type in any number of terms. By default, it is assumed that you want to search for all of the given terms within each of the textual units contained in the whole Index Thomisticus database, and that the given terms may appear in any order whatsoever with the only restriction that they be contiguous. However, you may override these settings by pressing the **options** button and modifying the defaults, or by using the advanced expression syntax given below. Having typed in the desired terms or expression, you can always press any of the following buttons:

concordances To read occurrences of the given terms (or expression) in their respective context.

terms To refine word selection for each given term by checking or unchecking individual forms, and to examine statistical information regarding their occurrence in the works that are selected.

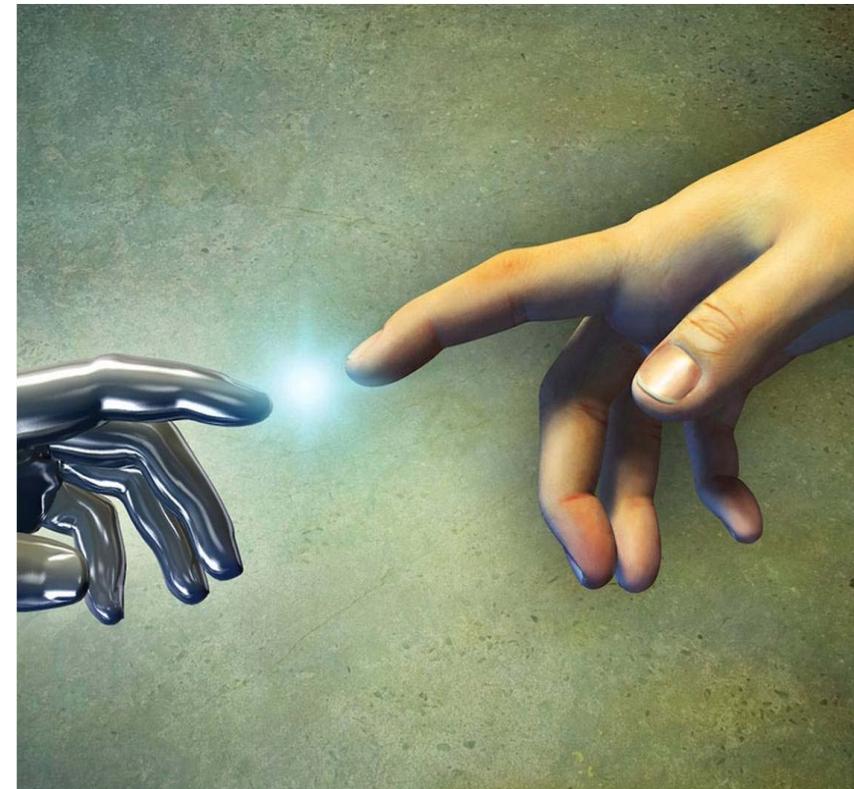
works To examine statistical information concerning the occurrences of the given terms or expression in all listed works.
Notice: You can press this button at any time to select which works (individually or by category) will be included in the search.

Terms:

Search terms may be expressed in three different manners: by supplying the desired word form, by supplying a lemma entry code (preceded by the # symbol) in order to select all of its forms, or by supplying a word form (preceded by the = symbol) that will serve as a sample for the selection of all of the forms belonging to the same lemma entry.

Table of Contents

- Check in
- The history of Digital Humanities: a brief look
- **Humanities research cultures & Cultural Heritage**
- Humanities research data
- Example I: Organizing Variations of Hieratic Script
- Example II: Letters of Brothers Grimm, Savigny & Bang
- Conclusion

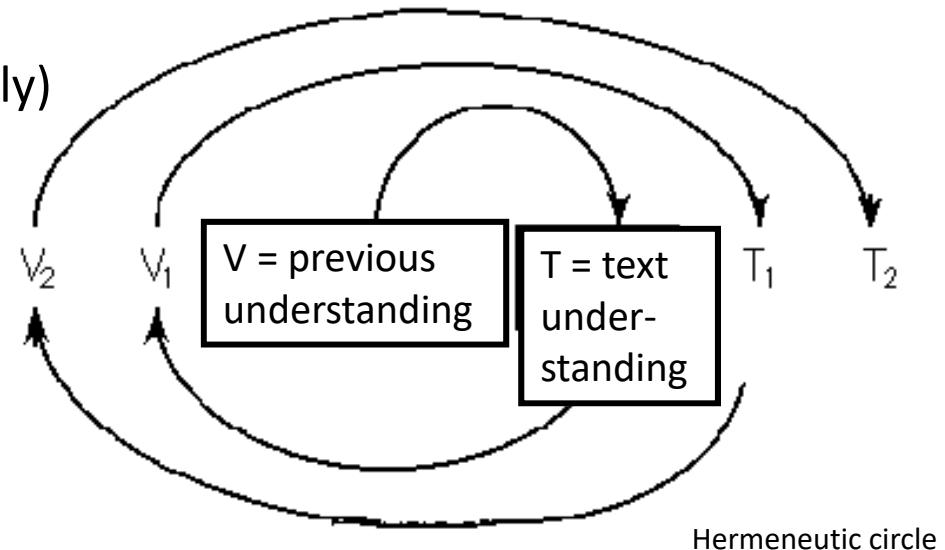


Humanities Research Cultures – the Hermeneutic Approach



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Acquisition and representation of cultural products of humankind (in the broadest sense)
 - Preserve & archive (but also discard, forget & destroy)
 - Copy, share, update & change
 - recording & disambiguate
 - Editing, commenting & explaining
 - Analyze & evaluate (also quantitatively)
 - Interlink & network
 - Interpret & understand
- A fundamentally unfinished and always to be renewed process



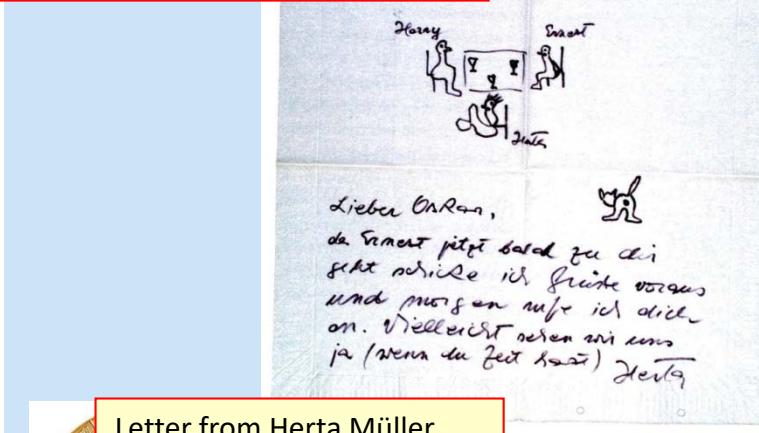
Diversity of Cultural Heritage



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Inka-Quipu, Museo Amano, Lima



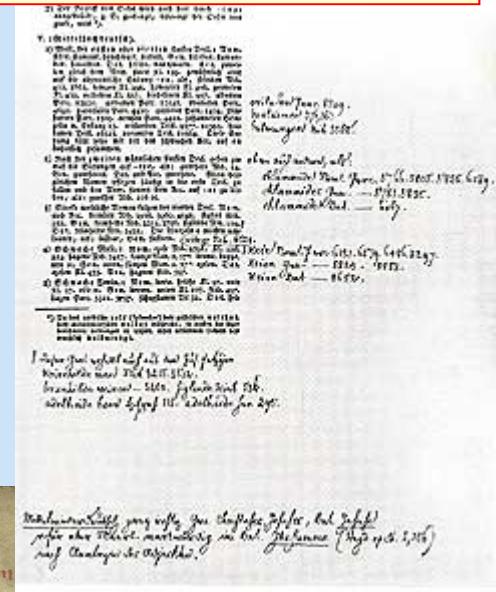
Letter from Herta Müller



Phaistos disc



slip boxes & Grimm's personal copy of the Grimms' Dictionary



Nibelungen C (BLB)



Fragility of Cultural Heritage



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Das digitale Historische Archiv Köln

Das DHAK wurde wenige Tage nach dem Einsturz des Historischen Archivs der Stadt Köln im März 2009 gegründet. Aus der als Nothilfe gedachten Idee entwickeln wir den digitalen, kollaborativen Zugang zum Archiv.



Erlernen Sie hier mehr zum Projekt

Blog Links Impressum Kontakt

Municipal Archive Cologne (Köln)

- 65.000 charters (since 922)
- 26 kilometers of records
- 104.000 maps & plans
- 50.000 posters
- 818 estates & collections

Collapse 3.3.2009

- 90% buried
- 85% recovered
- 200 conservators
- 30 years
- 370 mio. Euro
- only some individual items digitized

Fragility of Cultural Heritage



Ancient city of Palmyra

- destroyed october 2015 by IS
- reconstruction based on 3D records
- 3D-print of a (downscaled) replica 2016

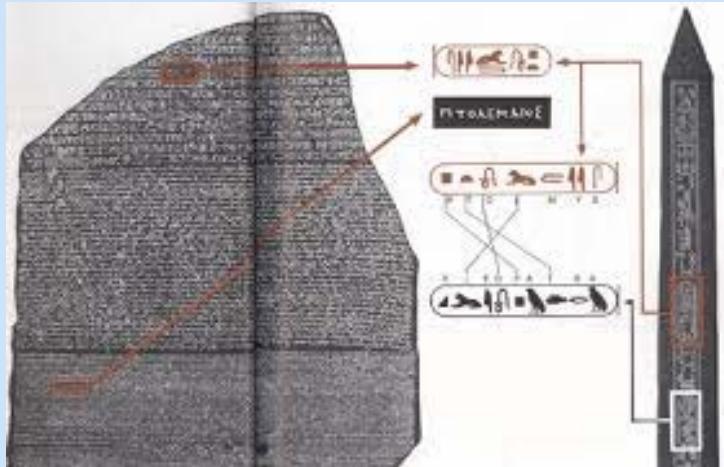
Buddha statues in Bamyan

- destroyed march 2001 by Taliban
- no „real“ reconstruction possible
- reconstruction based on 2D and 3D records
- hologram replica

Necessity of Context



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Rosetta Stone (196 BC)

- text in 3 languages and 3 script types: hieroglyphic-demotic-greek script
- cartouche with proper names = the key for (re-)deciphering

Phaistos Disc (1850-1550 BC?)

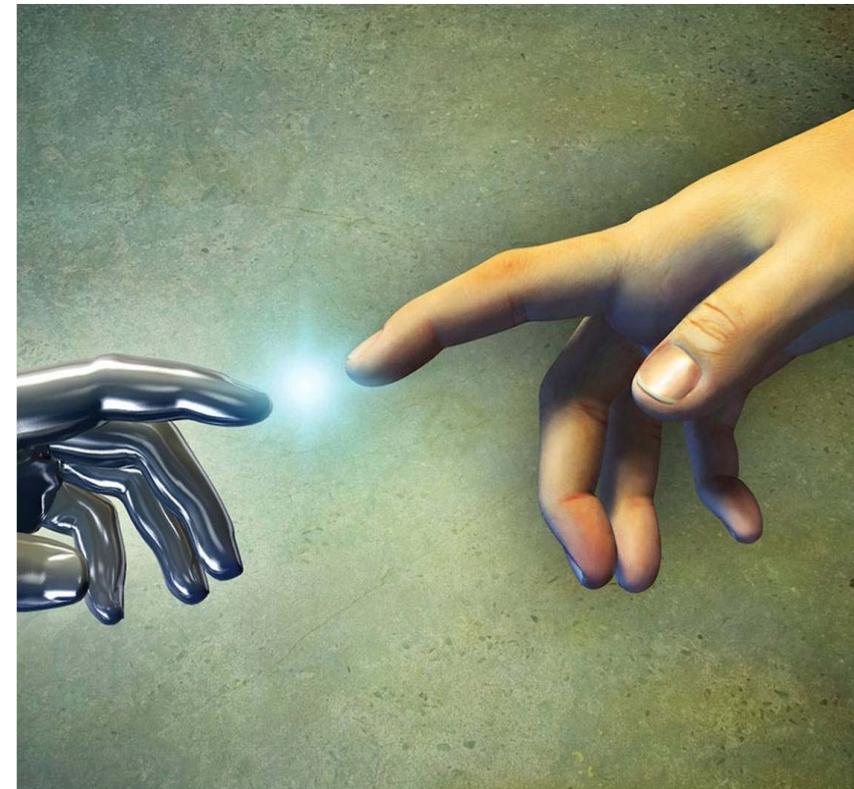
- individual item, unknown language & script
- 45 distinct single signs & text outline elements
- not decipherable due to lack of context

Table of Contents



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Check in
- The history of Digital Humanities: a brief look
- Humanities research cultures & Cultural Heritage
- **Humanities research data**
- Example I: Organizing Variations of Hieratic Script
- Example II: Letters of Brothers Grimm, Savigny & Bang
- Conclusion



Diversity of Cultural Heritage Data



TECHNISCHE
UNIVERSITÄT
DARMSTADT



DARIAH-DE-Stakeholdergremium
„Wissenschaftliche Sammlungen“: Oltersdorf,
Jenny und Schmunk, Stefan. 2016. „Von
Forschungsdaten und wissenschaftlichen
Sammlungen“. *Bibliothek Forschung und Praxis* 40
(2). <https://doi.org/10.1515/bfp-2016-0036>. S.181.

DARIAH-DE

“Data, without regard to their provenance, become digital research data in the context of a humanities and cultural research question, once it has been collected, described, evaluated and / or generated and stored in machine-readable form for the purpose of traceability of research results and for archiving, citability, and further processing.”

Diversity of Cultural Heritage Data



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Humanities research data: all digital data resulting from experiments, measurements, simulations, source references, transcriptions, editions, surveys, excavations, directories, recordings, etc.



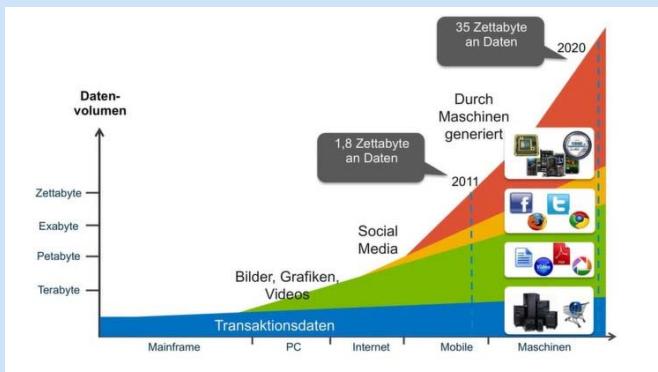
Formats, Media, Multimodalities

- ✓ texts, images, multi-dimensional models, audio, video, tables, databases, software (source code and applications)
- ✓ technical and device-specific raw data
- ✓ documentation
- ✓ software required for their understanding
- ✓ discipline-specific in different aggregation levels & mostly in different special digital formats

Big data – long data

Big data

Amounts of data that are too large, too complex, too fast-paced, or too weakly structured to be evaluated using manual and traditional methods of data processing.



Long data

Datasets of historical significance that provide not just a snapshot, but a comprehensive picture, drawing the evolution and the evolution of processes and interactions over long periods of time because long data provides context.

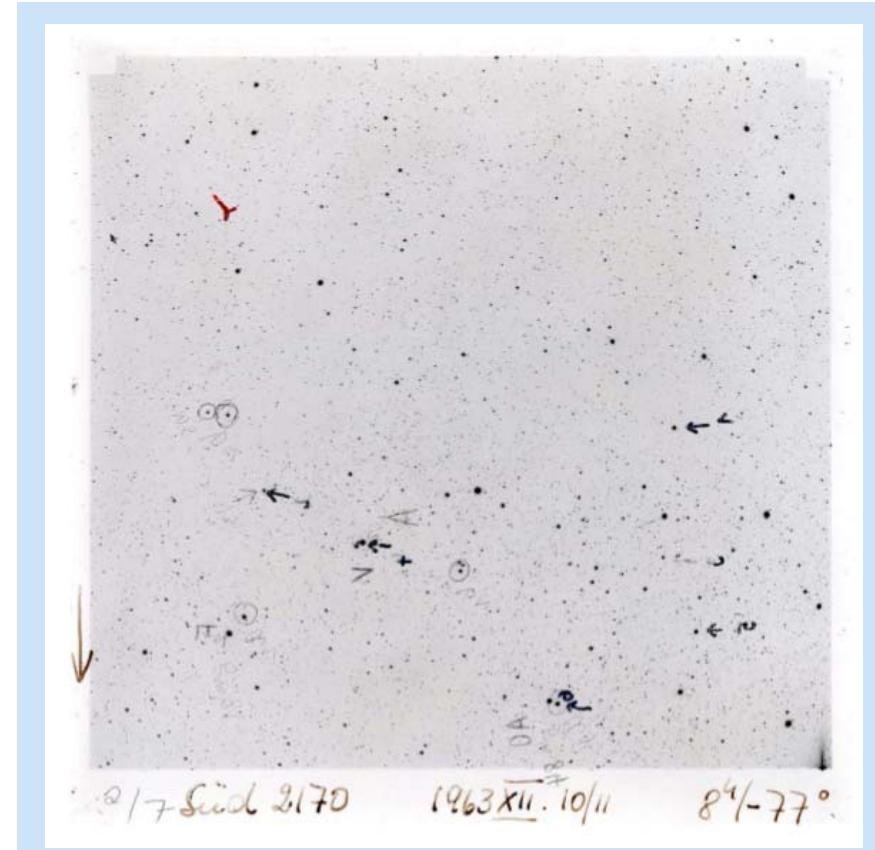


Long data



TECHNISCHE
UNIVERSITÄT
DARMSTADT

By “long data”, I mean datasets that have **massive historical sweep** — taking you from the dawn of civilization to the present day. (...) Because as beautiful as a snapshot is, how much richer is a **moving picture**, one that allows us to see how **processes and interactions** unfold over time? (...) — because long data provides **context**. (...) Want to see more clearly how knowledge is preserved? Use copies of a text created over a thousand years. (...)



Samuel Arbesman: Stop Hyped Big Data and Start Paying Attention to ‘Long Data’, 01.29.13 9:30 AM

<http://www.wired.com/opinion/2013/01/forget-big-data-think-long-data/>

<https://www.plate-archive.org/applause/>

Digital Transformation of Cultural Heritage



TECHNISCHE
UNIVERSITÄT
DARMSTADT

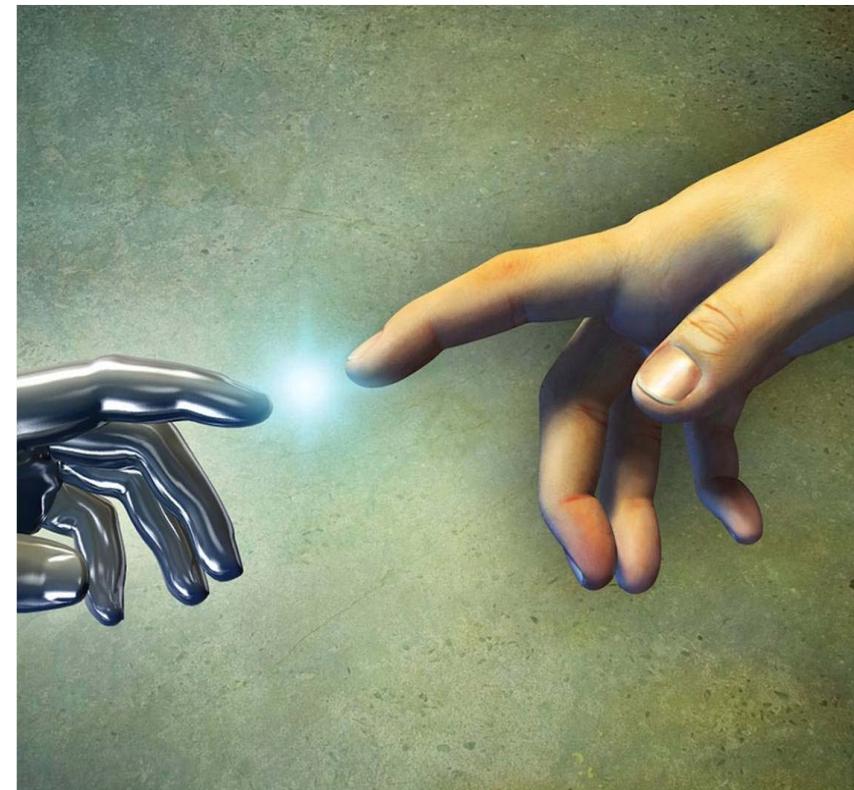
Humanities Collections

„[...] the **creation of digital objects** [...] is a crucial part of humanities research. [...] This is a fundamental difference between data-bases as they are used in the humanities and those that are used in the natural sciences. The **way** in which inscriptions are **photographed** or in which text corpora are **transcribed and encoded**, is **crucial for the way in which these research objects will be studied in the future.**“

Peursen, Wido T., van . “Editorial.” *Text Comparison and Digital Creativity*. Ed. Adriaan H. van der Weel, Ernst Thoutenhoofd, Wido T. van Peursen. Leiden: Brill, 2010. 11. Print.

Table of Contents

- Check in
- The history of Digital Humanities: a brief look
- Humanities research cultures & Cultural Heritage
- Humanities research data
- **Example I: Organizing Variations of Hieratic Script**
- Example II: Letters of Brothers Grimm, Savigny & Bang
- Conclusion



Example I: Organizing Variations of Hieratic Script



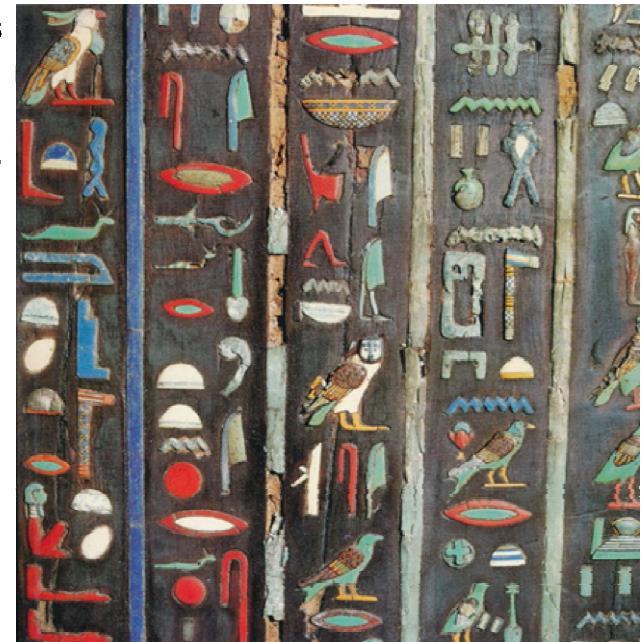
TECHNISCHE
UNIVERSITÄT
DARMSTADT

3000 years of written culture in Egypt: hieroglyphs



Sog. „Weiße Kapelle“
12. Dynastie
Pharao Sesotris I.
(ca. 1900 v. Chr.)

Sarg des
Djeddjehutiuafanch
Turin
(4. Jh. v. Chr.)



aus: Leclant, Ägypten III, Umschlag und Abb. 182

Foto: T. Konrad

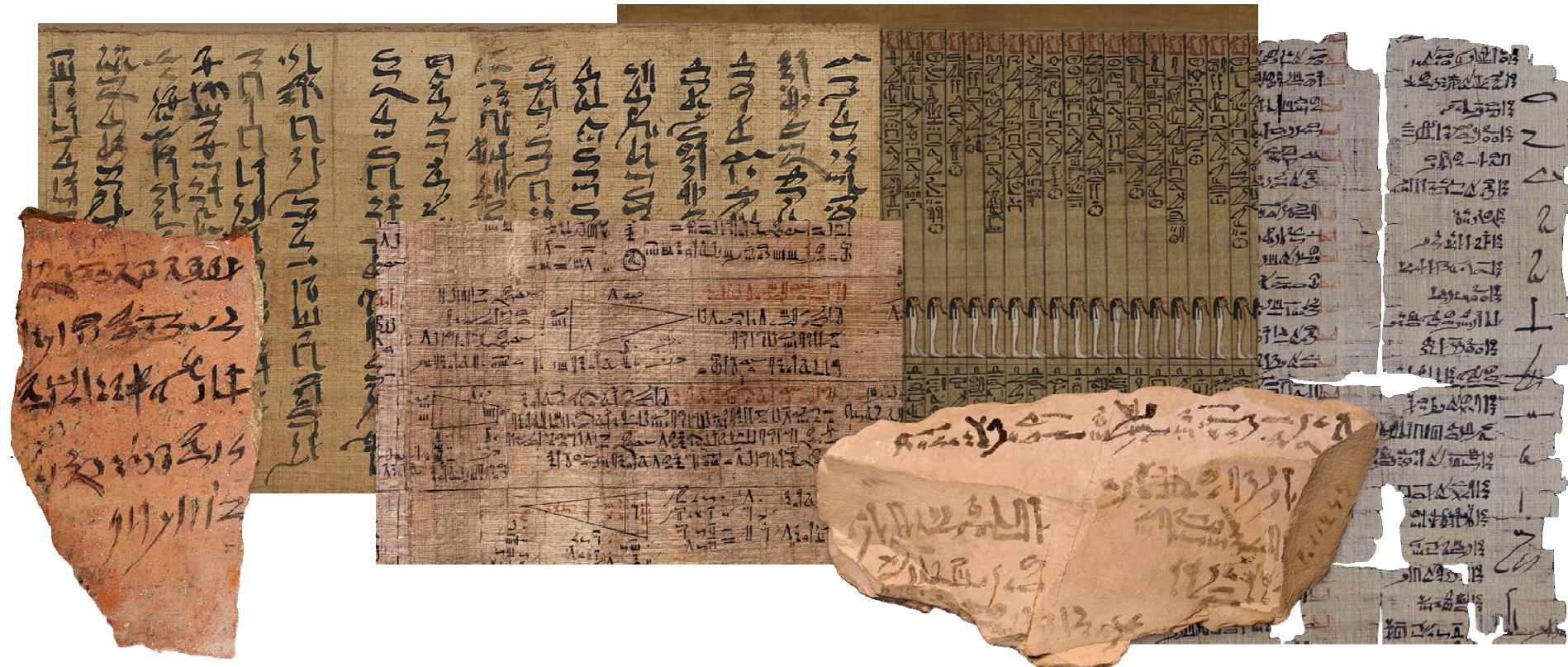
© AKU T. Konrad / S. Gülden + U. Verhoeven-van Elsbergen, K. von der Moezel, M. Leuk,
C. Krause, A. Rapp. <https://www.blogs.uni-mainz.de/fb07-aku-eng/>

Example I: Organizing Variations of Hieratic Script



TECHNISCHE
UNIVERSITÄT
DARMSTADT

3000 years of written culture in Egypt: hieratic script (hand written, ,italic' script)



Example I: Hieratic Scripts – Printed Paleographies



TECHNISCHE
UNIVERSITÄT
DARMSTADT

HIERATISCHE PALÄOGRAPHIE
DIE AEGYPTISCHE BUCHSCHRIFT
IN IHRER ENTWICKLUNG
BIS ZUM BEGINN DER ACHTEHENTEN DYNASTIE
DARGESTELLT
VON
GEORG MÖLLER
ERSTER BAND
BIS ZUM BEGINN DER ACHTEHENTEN DYNASTIE
MIT NEUN TAFELN SCHRIFTSTÜCKEN
ZWEITE, VERBESSERTE AUFLAGE.

Reference sign

Source

Date

196. **54,15.**
Basselbe Zeichen,
abge-
kürzte
Form.
1,2

A. Männer.

I 189 - 196 B

196. **54,15.**
Basselbe Zeichen,
abge-
kürzte
Form.
1,2

B

Miller, Paläographie 1

1

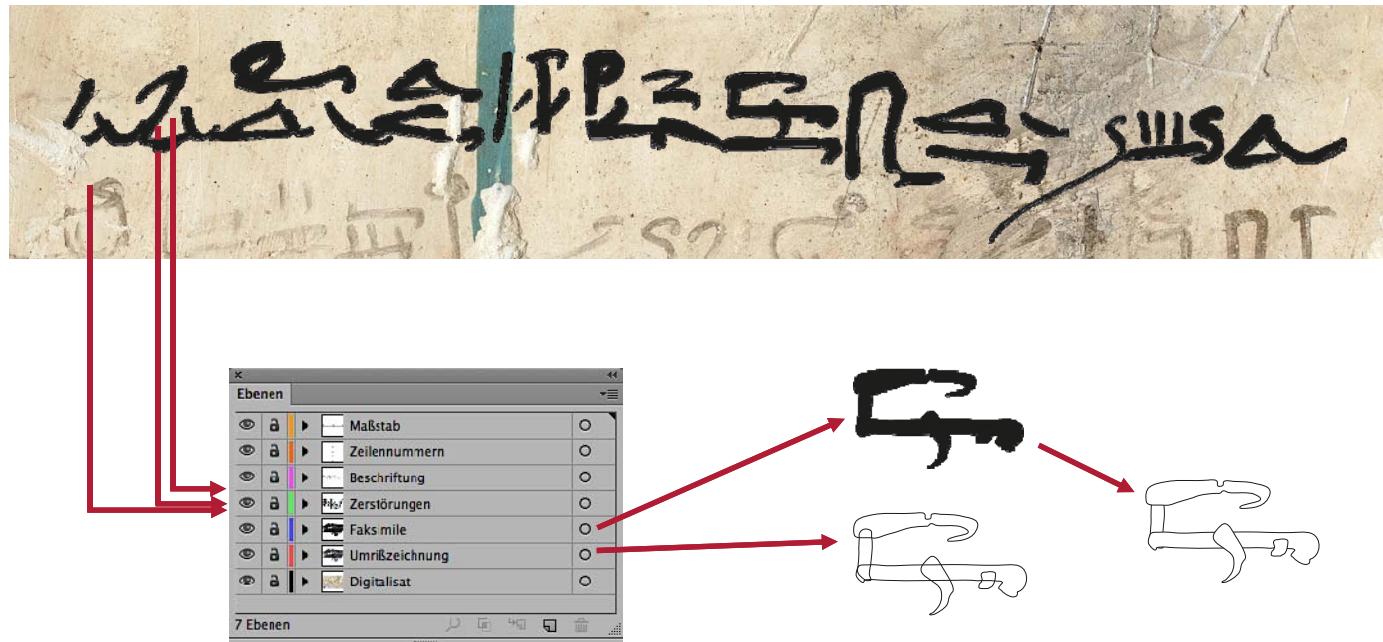
18

Example I: Hieratic Scripts – Drawing Digital Faksimiles



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Vector graphics of ‚hieratograms‘



Example I: Hieratic Scripts – Drawing Digital Faksimiles



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Different formats

The screenshot illustrates the process of digitizing a physical manuscript page containing hieratic script. The original image is shown at the top. Below it, a software interface displays a list of actions taken during the processing of the image. A red box highlights the 'ai in eps' and 'als svg und png speichern' options. To the right, three separate windows show the resulting files:

- *.eps**: Shows a low-resolution version of the script.
- *.svg**: Shows a higher-resolution version of the script.
- *.png**: Shows another version of the script.

The software interface includes a toolbar at the top and a list of actions on the left, such as 'Alles auswählen' and 'Speichern unter'.

Name	Größe
Gar01_Moe33_D_Asslut_N13_1_TN18_1a.eps	472 KB
Gar04_Moe82_82ba_D_Asslut_N13_1_TN18_1a.eps	481 KB
Gar04_Moe82_82ba_D_Asslut_N13_1_TN18_1b.eps	484 KB
Gar04_Moe82_82ba_D_Asslut_N13_1_TN18_1c.eps	479 KB
Gar04_Moe119_120_D_Asslut_N13_1_TN18_1d.eps	487 KB
Gar04_Moe119_120_D_Asslut_N13_1_TN18_1e.eps	481 KB
Gar01_Moe103b_D_Asslut_N13_1_TN18_1a.eps	492 KB
Gar01_Moe192b_D_Asslut_N13_1_TN18_1b.eps	482 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1a.eps	482 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1b.eps	468 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1c.eps	498 KB
Gar33_Moe331_D_Asslut_N13_1_TN18_1a.eps	479 KB
Gar33_Moe331_D_Asslut_N13_1_TN18_1b.eps	477 KB
Gar33_Moe331_D_Asslut_N13_1_TN18_1c.eps	490 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1a.eps	464 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1b.eps	479 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1c.eps	489 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1d.eps	489 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1e.eps	482 KB
Gar27_Moe2008_D_Asslut_N13_1_TN18_1a.eps	479 KB
Gar27_Moe2008_D_Asslut_N13_1_TN18_1b.eps	479 KB

Name	Größe
Gar01_Moe33_D_Asslut_N13_1_TN18_1a.svg	262 KB
Gar04_Moe82_82ba_D_Asslut_N13_1_TN18_1a.svg	261 KB
Gar04_Moe82_82ba_D_Asslut_N13_1_TN18_1b.svg	259 KB
Gar04_Moe82_82ba_D_Asslut_N13_1_TN18_1c.svg	253 KB
Gar04_Moe119_120_D_Asslut_N13_1_TN18_1d.svg	253 KB
Gar04_Moe119_120_D_Asslut_N13_1_TN18_1e.svg	258 KB
Gar01_Moe103b_D_Asslut_N13_1_TN18_1a.svg	259 KB
Gar01_Moe192b_D_Asslut_N13_1_TN18_1b.svg	254 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1a.svg	256 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1b.svg	252 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1c.svg	250 KB
Gar33_Moe331_D_Asslut_N13_1_TN18_1a.svg	262 KB
Gar33_Moe331_D_Asslut_N13_1_TN18_1b.svg	263 KB
Gar33_Moe331_D_Asslut_N13_1_TN18_1c.svg	265 KB
Gar33_Moe338_D_Asslut_N13_1_TN18_1a.svg	257 KB
Gar33_Moe338_D_Asslut_N13_1_TN18_1b.svg	253 KB
Gar33_Moe338_D_Asslut_N13_1_TN18_1c.svg	258 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1a.svg	259 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1b.svg	258 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1c.svg	258 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1d.svg	258 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1e.svg	258 KB
Gar27_Moe2008_D_Asslut_N13_1_TN18_1a.svg	258 KB
Gar27_Moe2008_D_Asslut_N13_1_TN18_1b.svg	258 KB

Name	Größe
Gar01_Moe33_D_Asslut_N13_1_TN18_1a.png	7 KB
Gar04_Moe82_82ba_D_Asslut_N13_1_TN18_1a.png	16 KB
Gar04_Moe82_82ba_D_Asslut_N13_1_TN18_1b.png	26 KB
Gar21_Moe119_120_D_Asslut_N13_1_TN18_1a.png	16 KB
Gar21_Moe119_120_D_Asslut_N13_1_TN18_1b.png	4 KB
Gar21_Moe119_120_D_Asslut_N13_1_TN18_1c.png	4 KB
Gar24_Moe119_120_D_Asslut_N13_1_TN18_1d.png	4 KB
Gar24_Moe119_120_D_Asslut_N13_1_TN18_1e.png	4 KB
Gar01_Moe103b_D_Asslut_N13_1_TN18_1a.png	25 KB
Gar01_Moe192b_D_Asslut_N13_1_TN18_1b.png	21 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1a.png	15 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1b.png	13 KB
Gar17_Moe282_283var_D_Asslut_N13_1_TN18_1c.png	22 KB
Gar33_Moe331_D_Asslut_N13_1_TN18_1a.png	17 KB
Gar33_Moe331_D_Asslut_N13_1_TN18_1b.png	7 KB
Gar33_Moe338_D_Asslut_N13_1_TN18_1a.png	11 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1a.png	20 KB
Gar12_Moe469var_D_Asslut_N13_1_TN18_1b.png	10 KB
Gar3_Moe537_D_Asslut_N13_1_TN18_1a.png	28 KB
Gar5_Moe540_D_Asslut_N13_1_TN18_1a.png	23 KB
Gar27_Moe2008_D_Asslut_N13_1_TN18_1a.png	8 KB
Gar27_Moe2008_D_Asslut_N13_1_TN18_1b.png	19 KB

Example I: Hieratic signs in XML-Code



Charactersize (Pixel x 2.54 cm) / 72 ppi = cm

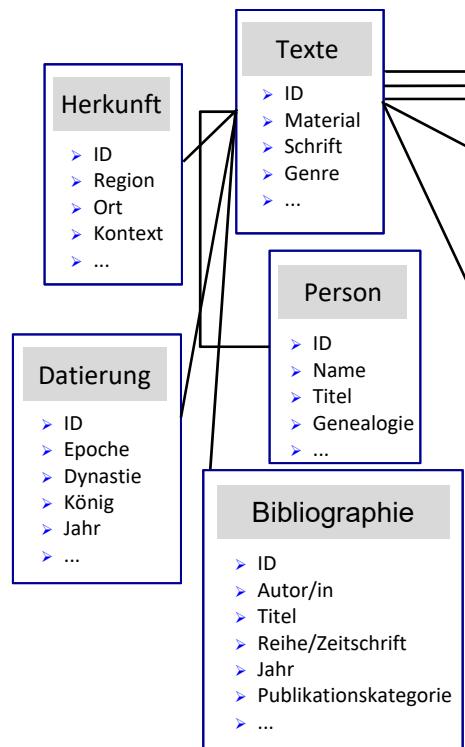


```
<?xml version="1.0" encoding="utf-8"?>
<!-- Generator: Adobe Illustrator 16.0.0, SVG Export Plug-In . SVG Version: 6.00 Build 0) --&gt;
&lt;!DOCTYPE svg PUBLIC "-//W3C//DTD SVG 1.1//EN" "http://www.w3.org/Graphics/SVG/1.1/DTD/svg11.dtd"&gt;
&lt;svg version="1.1" id="Ebene_1" xmlns:x=&amp;ns_extend;" xmlns:i=&amp;ns_ai;" xmlns:graph=&amp;ns_graphs;" 
      xmlns="http://www.w3.org/2000/svg" xmlns:xlink="http://www.w3.org/1999/xlink" x="0px" y="0px"
<b>width="50.043px" height="29.552px" viewBox="0 0 50.043 29.552" enable-background="new 0 0 50.043 29.552"
xml:space="preserve">
<switch>
    <foreignObject requiredExtensions=&ns_ai;" x="0" y="0" width="1" height="1">
        <i:pgfRef xlink:href="#adobe_illustrator_pgf">
            </i:pgfRef>
    </foreignObject>
    <g i:extraneous="self">
        <path fill="#1E1E1C" stroke="#1B181C" stroke-width="0.15" d="M5.452,10.883c0.08-
1.092,0.088-2.185-0.088-3.267
        [...]
        <path fill="..."/>
        <path fill="..."/>
    </g>
</switch>
```

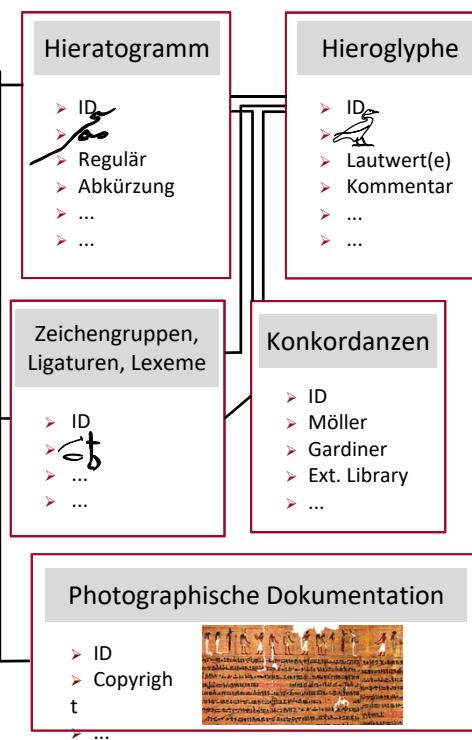
Example I: Databases



Metadata database



Media database



Example I: Browsing & Searching Script Variations



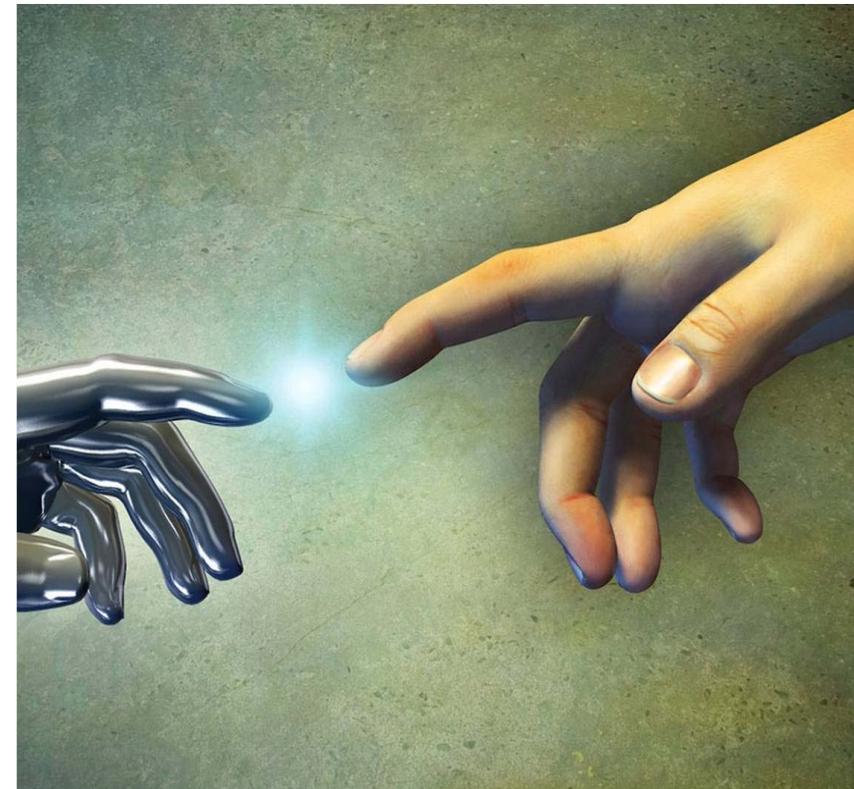
TECHNISCHE
UNIVERSITÄT
DARMSTADT

Hieroglyphe									
Berlin P. 3073 (von -305 bis -30)	London EA 10188 [palimpsest new] (-311)	Berlin P. 3057 (P. Schmitt) (von -332 bis -30)	Berlin P. 13242 Ro (von -30 bis 14)	Leiden (Leemans) T 32 (65)	London EA 10672 (von -30 bis 283)	Berlin P. 7809 (von 101 bis 150)	Edinburgh A 1956.313 (P. Rhind 1) (-9)	Edinburgh A 1956.314 (P. Rhind 2) (-11)	

Design: T. Yousef

Table of Contents

- Check in
- The history of Digital Humanities: a brief look
- Humanities research cultures & Cultural Heritage
- Humanities research data
- Example I: Organizing Variations of Hieratic Script
- **Example II: Letters of Brothers Grimm, Savigny & Bang**
- Conclusion

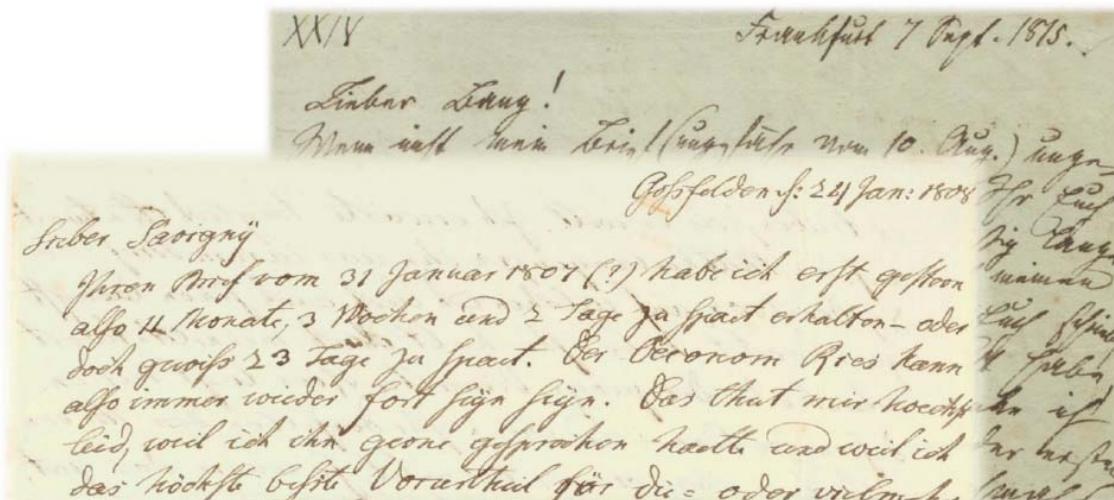


Example II: Letters of Brothers Grimm, Savigny & Bang



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Jacob Grimm (1785-1863), founder of German philology and ancient studies
- Wilhelm Grimm (1786-1859), philologist
- Friedrich Carl von Savigny (1779-1861), founder of historical legal studies
- Johann Heinrich Christian Bang (1774-1851), parish priest, pedagogue, farmer



Editors:
Sabine Bartsch
Mareike Bassenge
Luise Borek
Franziska Horn
Rotraut Fischer
Andrea Rapp

<https://www.linglit.tu-darmstadt.de/index.php?id=2834>

Example II: Letters of Brothers Grimm, Savigny & Bang



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Virtual Research environment TextGrid

The screenshot shows the TextGrid software interface. On the left is a Navigator panel listing various files, including 'HA_Autographen_K50' and 'Ms-725-44_53_0002'. The central area displays a handwritten letter in German. The right side is a Metadaten-Editor window for 'Ms-725-44_53_0002'. The editor contains fields for 'Title(s)', 'Identifier(s)', 'Format', and 'Rights Holder(s)'. The 'Title(s)' field contains 'Ms-725-44_53_0002'. The 'Identifier(s)' section lists '340 Bang I Nr. 3_1', '340 Bang I Nr. 3_2', '340 Bang I Nr. 3_3', '340 Bang I Nr. 3_4', and 'Ms-725-44_53_0001'. The 'Format' field is set to 'image/tiff'. The 'Rights Holder(s)' section lists 'Universitätsbibliothek Marburg' as the name and 'Identifier (xs:anyURI)'.

Navigator

Ms-725-44_53_0002

Metadaten-Editor

Title(s)

Title *

Ms-725-44_53_0002

Entfernen

HinzufügenTitle

Identifier(s)

Format *

image/tiff

Rights Holder(s)

Name

Universitätsbibliothek Marburg

Identifier (xs:anyURI)

Is corporate body?

Entfernen

HinzufügenRights Holder

Humanities Metadata Standards



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Guidelines of the Text Encoding Initiative TEI
- A Consortium that discusses and develops standards for different text / document types in an open and collaborative way

- TEI Guidelines: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/>

- TEI by Example: <http://tbe.kantl.be/TBE/TBE.htm>

- Oxford Teaching Pages: <http://tei.oucs.ox.ac.uk/Talks/>

TEI-Modules



Module name	Formal public identifier	Where defined	
analysis	Analysis and Interpretation	17 Simple Analytic Mechanisms	
certainty	Certainty and Uncertainty	21 Certainty, Precision, and Responsibility	
core	Common Core	3 Elements Available in All TEI Documents	
corpus	Metadata for Language Corpora	15 Language Corpora	
dictionaries	Print Dictionaries	9 Dictionaries	
drama	Performance Texts	7 Performance Texts	
figures	Tables, Formulae, Figures	14 Tables, Formulae, and Graphics	
gaiji	Character and Glyph Documentation	5 Representation of Non-standard Characters and Glyphs	
header	Common Metadata	2 The TEI Header	
iso-fs	Feature Structures	18 Feature Structures	
linking	Linking, Segmentation, and Alignment	16 Linking, Segmentation, and Alignment	
msdescription	Manuscript Description	10 Manuscript Description	
namesdates	Names, Dates, People, and Places	13 Names, Dates, People, and Places	
nets	Graphs, Networks, and Trees	19 Graphs, Networks, and Trees	
spoken	Transcribed Speech	8 Transcriptions of Speech	
tagdocs	Documentation Elements	22 Documentation Elements	
tei	TEI Infrastructure	1 The TEI Infrastructure	http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ST.html#STMA
textcrit	Text Criticism	12 Critical Apparatus	
textstructure	Default Text Structure	4 Default Text Structure	
transcr	Transcription of Primary Sources	11 Representation of Primary Sources	
verse	Verse	6 Verse	

Standards for Names, Dates, People, and Places



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Encoding persons, places, organizations, time and date ...
- Can be stored as lists in the TEI-Header (or in separate files)
- Authority files (examples):
 - OGND ('Gemeinsame Normdatei') <http://swb.bsz-bw.de>
 - Catalogue of the German National Library (DNB) <https://portal.dnb.de>
 - ISO online browsing platform <https://www.iso.org/obp/ui/#search>
 - Language subtag lookup <https://r12a.github.io/app-subtags/>
 - Getty Thesaurus of Geographic Names®
<http://www.getty.edu/research/tools/vocabularies/tgn/>
 - FOAF (Friend of a Friend) <http://www.foaf-search.net>

Example II: Letters of Brothers Grimm, Savigny & Bang



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Metadata for physical object 'letter'

The diagram illustrates the mapping between a user interface for describing a physical object and its corresponding XML schema. A blue arrow points from the user interface on the left to the XML code on the right.

User Interface (Left):

- Section:** 2. Beschreibung des materiellen Zustands
- Buttons:** neue Beschreibung des materiellen Zustands
- Text:** Art und Qualität des Beschreibstoffes: zartgrünes Papier
- Text:** Stempel: Stempel auf der Rückseite ▷ UB Mbg. ◁ handschriftlicher Vermerk: ▷ Hs 725:52 ◁
- Text:** Blattanzahl: 1 Blatttyp: double_leaf
- Text:** Maße der Blätter in cm:
 - Höhe: 16,4
 - Breite: 12,6
- Text:** Einschätzung des materiellen Zustandes des Briefes:
gut erhalten
- Text:** Beschreibung der Gestaltung:
beidseitig beschrieben

XML Schema (Right):

```
<physDesc>
  <objectDesc>
    <supportDesc>
      <support>
        <material>zartgrünes Papier</material>
        <stamp>Stempel auf der Rückseite
          <mentioned>UB Mbg.</mentioned>handschriftlicher
          Vermerk: <mentioned>Hs 725:52</mentioned></stamp>
        </support>
      <extent>
        <measure unit="double_leaf" quantity="1"></measure>
        <dimensions unit="cm">
          <height>16,4</height>
          <width>12,6</width>
        </dimensions>
      </extent>
      <condition>
        <p>gut erhalten</p>
      </condition>
    </supportDesc>
    <layoutDesc>
      <layout>
        <p>beidseitig beschrieben</p>
      </layout>
    </layoutDesc>
```

Example II: Letters of Brothers Grimm, Savigny & Bang



TECHNISCHE
UNIVERSITÄT
DARMSTADT

XML transcript expatiating implicit informations & GND-IDs for persons

Wenn es Ihnen möglich wäre, lieber Bang, heute oder morgen herein zu kommen, So würden Sie mir dadurch einen großen Gefallen thun: ich habe einen <gestr.: B> wichtigen Brief von Br.² den ich den Sonntag beantworten muß, aber nicht ehe beantworten kann als bis ich Sie geßprochen habe. Mich hindert das Pockenfieber, zu Ihnen zu kommen. Auch der Magiſter³ wünscht Sie wegen ſeiner Sache zu ſprechen.

Ihr
Sav

² Clemens Wenzeslaus Maria Brentano (1778-1842) an Friedrich Carl von

Nürnberg, Germanisches Nationalmuseum, Ms. XXXVII: Friedrich Carl von Savigny an Johann Heinrich Christian Bang, ohne Ort, ohne Datum.



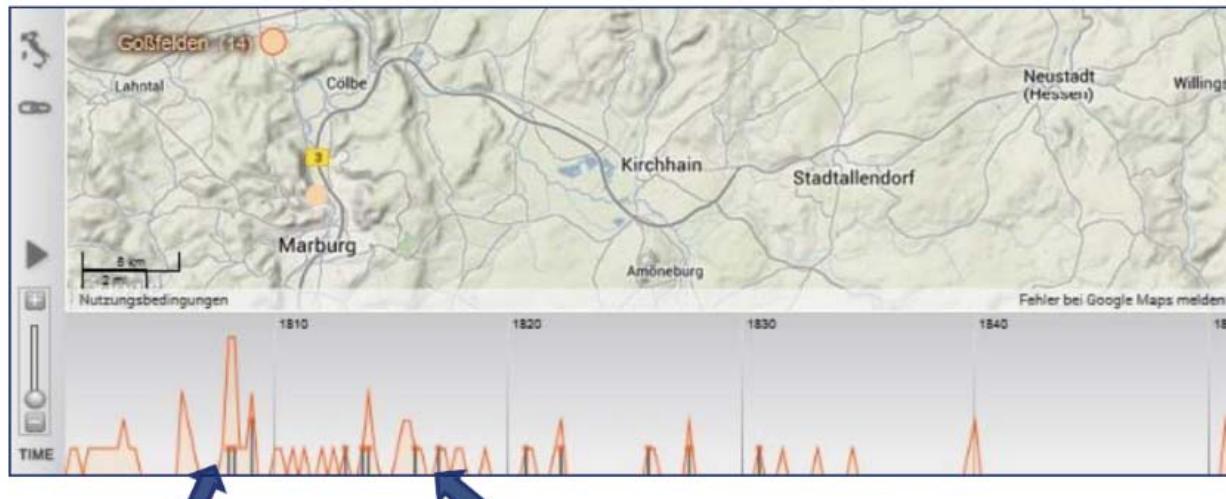
```
<p> Wenn es Ihnen möglich wäre, lieber Bang, heute oder morgen herein zu kommen, So würden Sie mir dadurch einen großen Gefallen thun: ich habe einen <del hand="author" rend="overstrike">B</del> wichtigen Brief von <abbr>Br. <expan>Brentano</expan> </abbr> <note type="commentary" resp="RF" xml:id="c_2"> <persName key="118515055">Clemens</persName> Wenzeslaus Maria Brentano (<date>1778-1842</date>)
```

Example II: Letters of Brothers Grimm, Savigny & Bang



TECHNISCHE
UNIVERSITÄT
DARMSTADT

DARIAH-Geobrowser: space-time-event visualization



```
<correspDesc>
  <sender>
    <persName key="116048778">"Bang,
    Johann Heinrich Christian"</persName>
  </sender>
  <addressee>
    <persName key="118605909">"Savigny,
    Friedrich Carl von"</persName>
  </addressee>
  <placeSender>
    <placeName>Goßfelden</placeName>
  </placeSender>
  <dateSender>
    <date when="1808-1-24">24.1.1808</date>
  </dateSender>
</correspDesc>
```

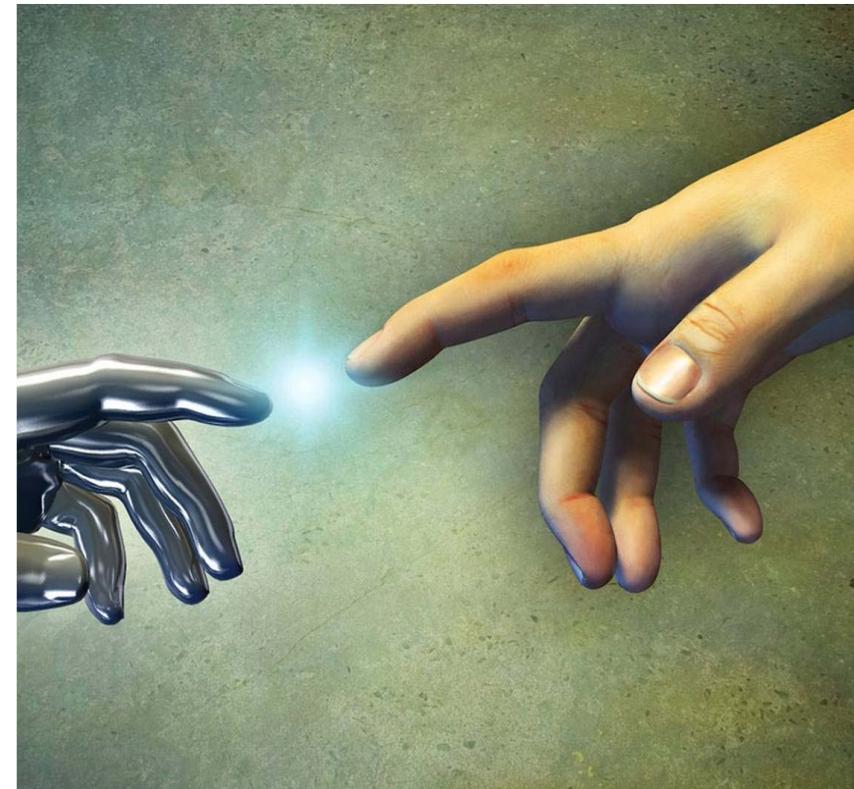
```
<correspDesc>
  <sender>
    <persName key="116048778">"Bang,
    Johann Heinrich Christian"</persName>
  </sender>
  <addressee>
    <persName key="118605909">"Savigny,
    Friedrich Carl von"</persName>
  </addressee>
  <placeSender>
    <placeName>Goßfelden</placeName>
  </placeSender>
  <dateSender>
    <date when="1817-03-12">12.3.1817</date>
  </dateSender>
</correspDesc>
```

Table of Contents



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Check in
- The history of Digital Humanities: a brief look
- Humanities research cultures & Cultural Heritage
- Humanities research data
- Example I: Organizing Variations of Hieratic Script
- Example II: Letters of Brothers Grimm, Savigny & Bang
- Conclusion



Conclusion



TECHNISCHE
UNIVERSITÄT
DARMSTADT

“All data is historical data: the product of a time, place, political, economic, technical, & social climate. If you are not considering why your data exists, and other data sets don’t, you are doing data science wrong”.

- We need contextual data not only to save and conserve the cultural (and every) knowledge but also for all other (scientific) purposes.

- The challenges are the same in the different research cultures: in the (natural) sciences, engineering sciences, social sciences, arts and humanities ...

The Turing
@turinginst

“All data is historical data: the product of a time, place, political, economic, technical, & social climate. If you are not considering why your data exists, and other data sets don’t, you are doing data science wrong”. -
@melissaterras
#TuringLectures
youtube.com/watch?v=bf6DsJ...
Translate Tweet



4:17pm · 5 Mar 2019 · Twitter Web Client



A large word cloud centered around the words "thank you" in various languages. The words are rendered in different colors and sizes, creating a dense, circular pattern. The most prominent words are "thank" and "you", which are repeated multiple times in various colors like blue, red, green, and yellow. Other visible words include "danke" (German), "gracias" (Spanish), "merci" (French), "多谢" (Chinese), "谢谢" (Chinese), "спасибо" (Russian), " obrigado" (Portuguese), "dank je" (Dutch), "teşekkür ederim" (Turkish), "go raibh maith agat" (Irish), "dziekuje" (Polish), "dakujem" (Czech), "arigato" (Japanese), "감사합니다" (Korean), "merci" (French), "ngiyabonga" (Swahili), "terima kasih" (Indonesian), "sukriya" (Malay), "kop khun krap" (Thai), "grazie" (Italian), and "dakujem" (Slovene).