



East Asian
Studies

RUHR-UNIVERSITÄT BOCHUM

FAKULTÄT FÜR OSTASIENWISSENSCHAFTEN | RELIGIONEN OSTASIENS

COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN
UND IHREN NACHBARDISZIPLINEN [BLOCK 1]

Lena Liefke



THEMENSCHWERPUNKTE

- Visualisierungstechniken und – Projekte kennenlernen
- Ein ausgewähltes Tool ausprobieren (*Voyant Tools*)

ZIELSETZUNG

- unterschiedlichste Formen der Visualisierung kennenlernen
- einen Eindruck von dem Stand der Technik / Forschung gewinnen
- Inspiration für die weitere Beschäftigung mit einem der Tools entwickeln
- einen ersten Schritt Richtung Korpusanalyse machen



ABLAUF

10.15-12.00Uhr (Visualisierungstechniken und Beispiele kennenlernen)

- **Einführung:** Möglichkeiten der Visualisierung
- **Recherche:** Beschäftigung mit einem Visualisierungstool (*VOSviewer, Gephi, ctext-Texttools*)

30min Pause

12.30-13.45Uhr (Beispiele aufbereiten und präsentieren)

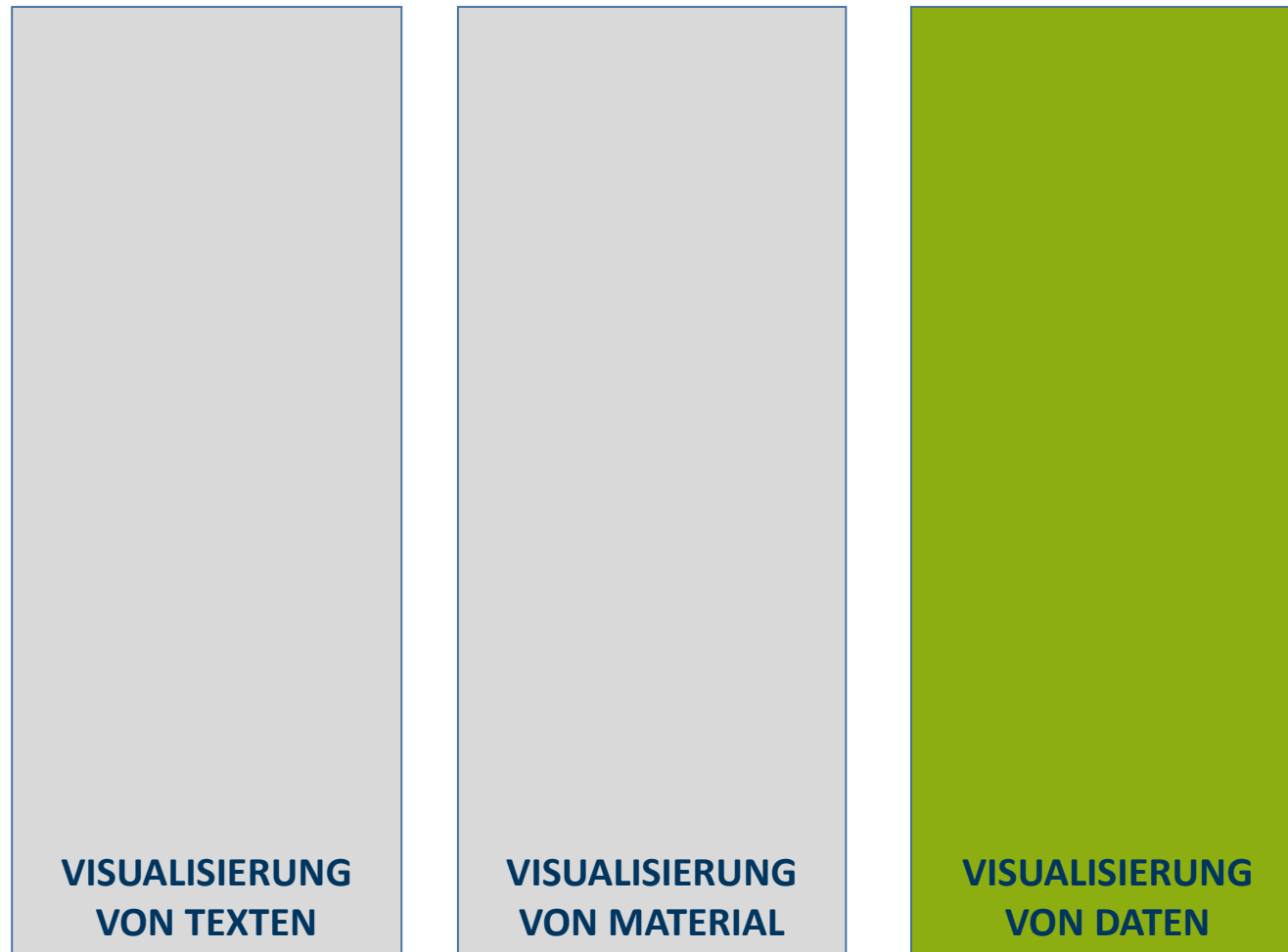
- **Aufbereitung:** Visualisierungsprojekt verstehen und aufbereiten
- **Präsentation:** Visualisierungsprojekt vorstellen

30 min Pause

14.15-15.45Uhr (Selbst Visualisierungen erstellen mit **Voyant Tools**)

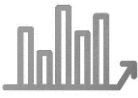
- **Vorstellung:** Funktionen von *Voyant Tools*
- **Anwendung:** Visualisierungen mit *Voyant Tools*

VISUALISIERUNG GEGENSTAND DER VISUALISIERUNG





Quelle: Jengolbeck, <https://www.youtube.com/watch?v=t0B2sxlumh0>



- **Information visualization**

- **Scientific visualization**

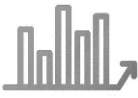
- **Infographics**

- **Static visualization**

- **Dynamic visualization**

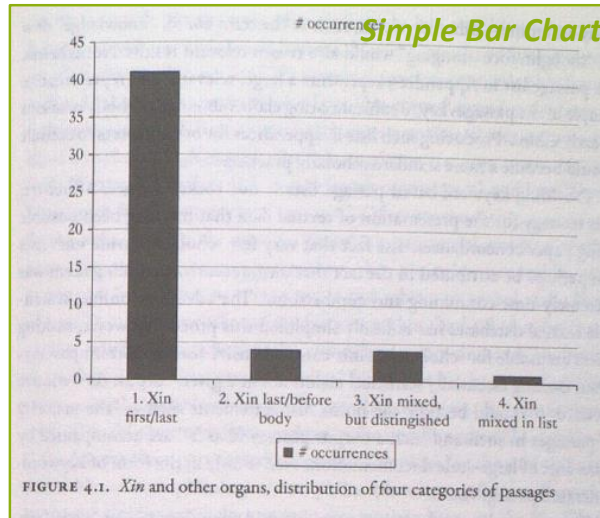
- **Interactive visualization**

**VISUALISIERUNG
VON DATEN**



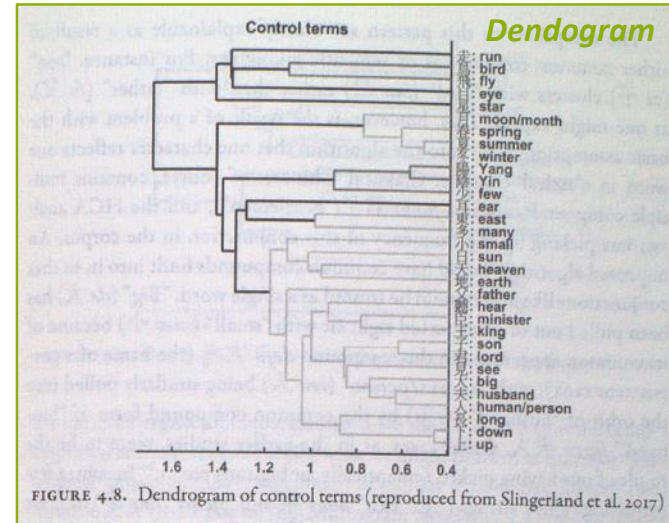
Static visualizations

Simple Bar Chart



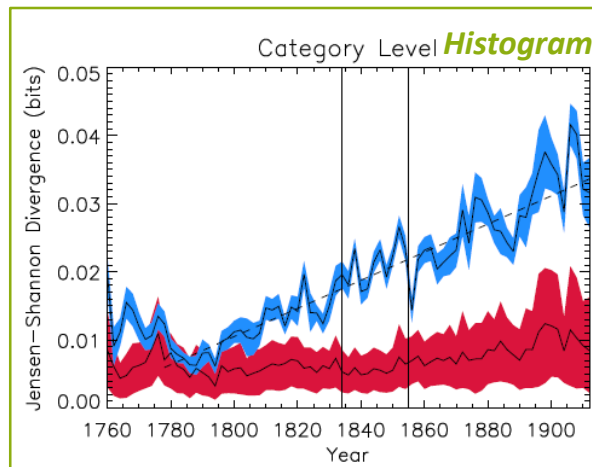
Slingerland (2019, p.149)

Dendrogram



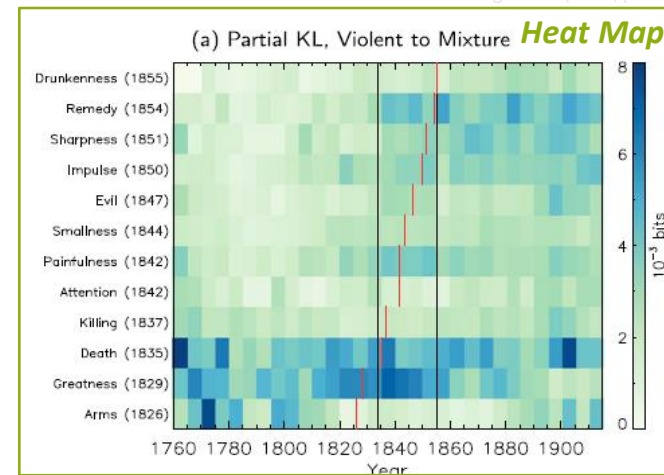
Slingerland (2019, p.185)

Histogram



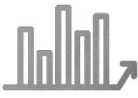
Klingenstein et al. (2014, p.9421)

Heat Map

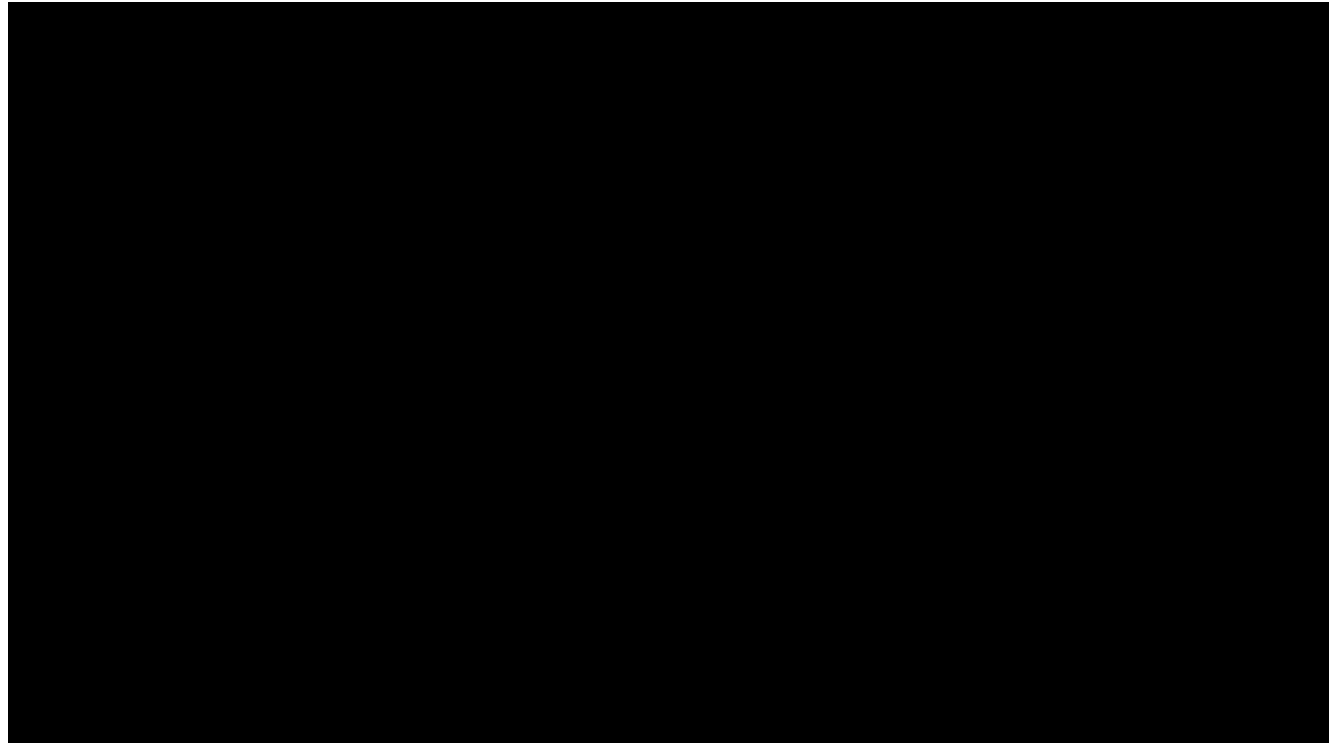


Klingenstein et al. (2014, p.9423)

VISUALISIERUNG
VON DATEN

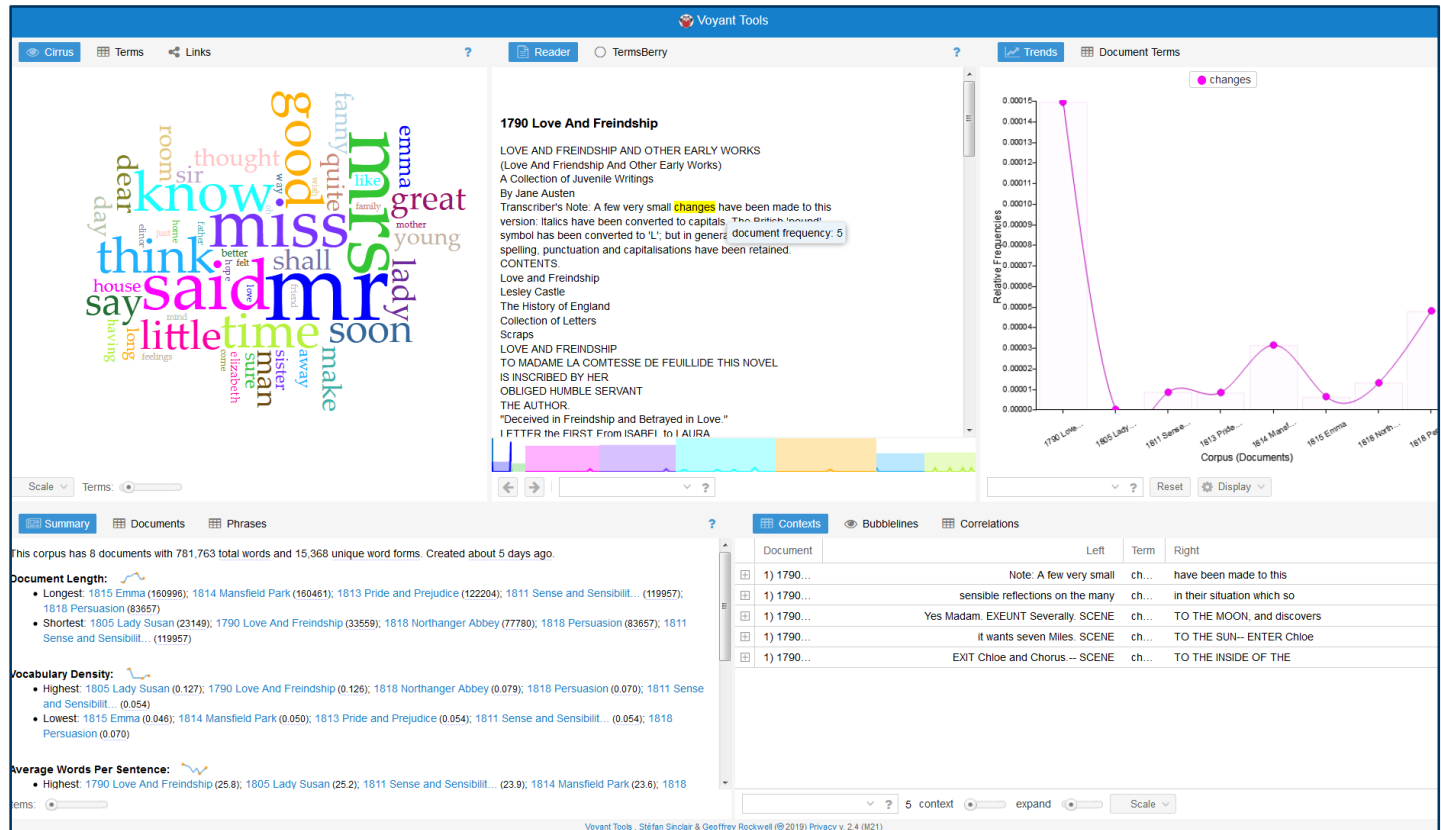
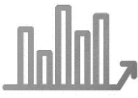


Dynamic visualizations



China Biographical Database Project: <https://projects.iq.harvard.edu/cbdb/home>

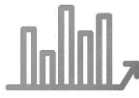
**VISUALISIERUNG
VON DATEN**



<https://voyant-tools.org/?corpus=austen>

VISUALISIERUNG
VON DATEN

VISUALIZATION / DIGITIZATION OF TEXTS



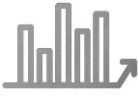
Ten Thousand Rooms Project

Digitale Bereitstellung (v.a. vormoderne) chinesischer Texte mit Möglichkeit der kollaborativen Transkription, Übersetzung, und Annotation

The screenshot displays the Ten Thousand Rooms Project interface. On the left, a table of contents lists chapters from 'front matter' to '17: 萬言 (chap. 14b)'. The central panel shows a digital scroll of the 'Kongzi jiyu' (Confucius's Sayings) text, with a specific section highlighted. On the right, an 'Annotations' panel is visible, showing a 'Translation' of the selected text. The translation discusses the importance of learning and the proper appearance of a nobleman, citing Kongzi's son Boyu.

Quelle: <https://tenthoutherooms.yale.edu/node/311/mirador?canvas=14116>

VISUALISIERUNG
VON TEXTEN



SAT

Digitale Bereitstellung der japanischen Edition des chinesischen buddhistischen Kanons mit Möglichkeit des Einsehens von gedruckten Editionen

The screenshot displays the SAT (Shinshu Arhanta Text) digital interface. The main window shows a Japanese Buddhist text, likely from the Shinshu Arhanta Text. The interface includes a sidebar on the left with navigation options like 'はじめに', '検索結果', '版面画像', '日本語訳', '英語訳', and 'IIIIF Linking'. The main text area is titled '註華嚴法界觀門' (Shinshu Arhanta Text) and shows a page number 'p. 683'. The right sidebar contains a table of contents or a list of related texts, including '註華嚴法界觀門 (No. 1884 宗密註) in Vol. 45'. The bottom navigation bar shows a series of thumbnails for pages 691, 690, 689, 688, 687, 686, 685, 684, and 683, with the current page (683) highlighted.

Quelle: <http://21dzk.l.u-tokyo.ac.jp/SAT2018/master30.php>

VISUALISIERUNG
VON TEXTEN



CBETA

Digitale Bereitstellung des chinesischen buddhistischen Kanons mit Suchfunktion, integriertem Wörterbuch und Anzeige der Häufigkeitsverteilung des Suchbegriffs

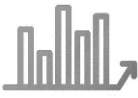


Quelle: http://cbetaonline.dila.edu.tw/zh/T1867_001

VISUALISIERUNG
VON DATEN

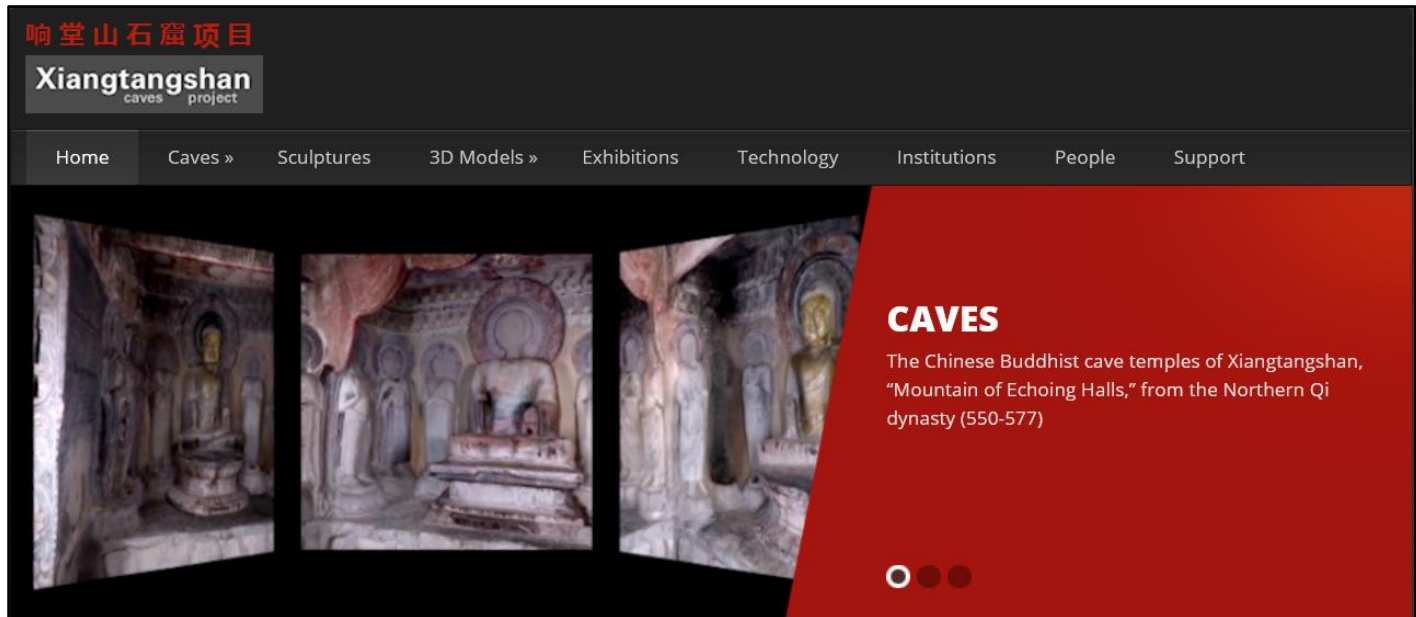
VISUALISIERUNG
VON TEXTEN

VISUALIZATION / DIGITIZATION OF MATERIAL



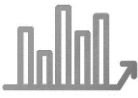
Xiangtangshan Caves

Rekonstruktion und virtuelle Animation (teils zerstörter) buddhistischer Grotten und dort (ehemals) beinhalteter Skulpturen



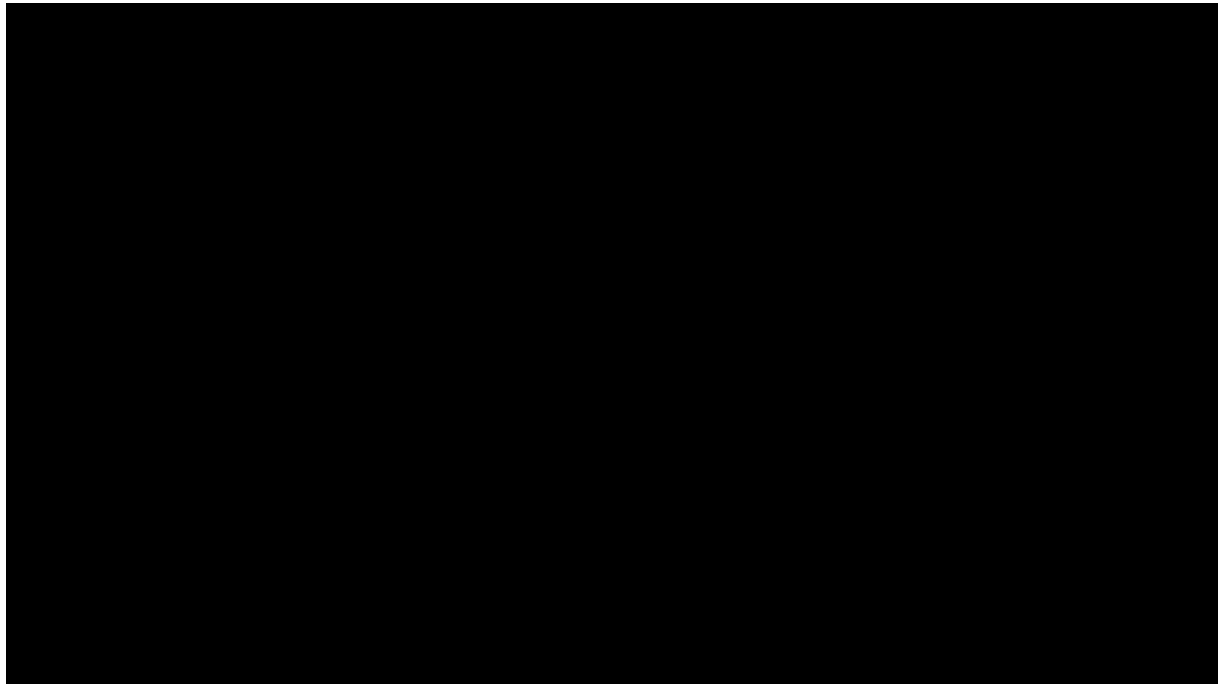
Quelle: <https://xts.uchicago.edu/>

VISUALISIERUNG
VON MATERIAL



Rome Reborn

Rekonstruktion und virtuelle Animation des alten Roms



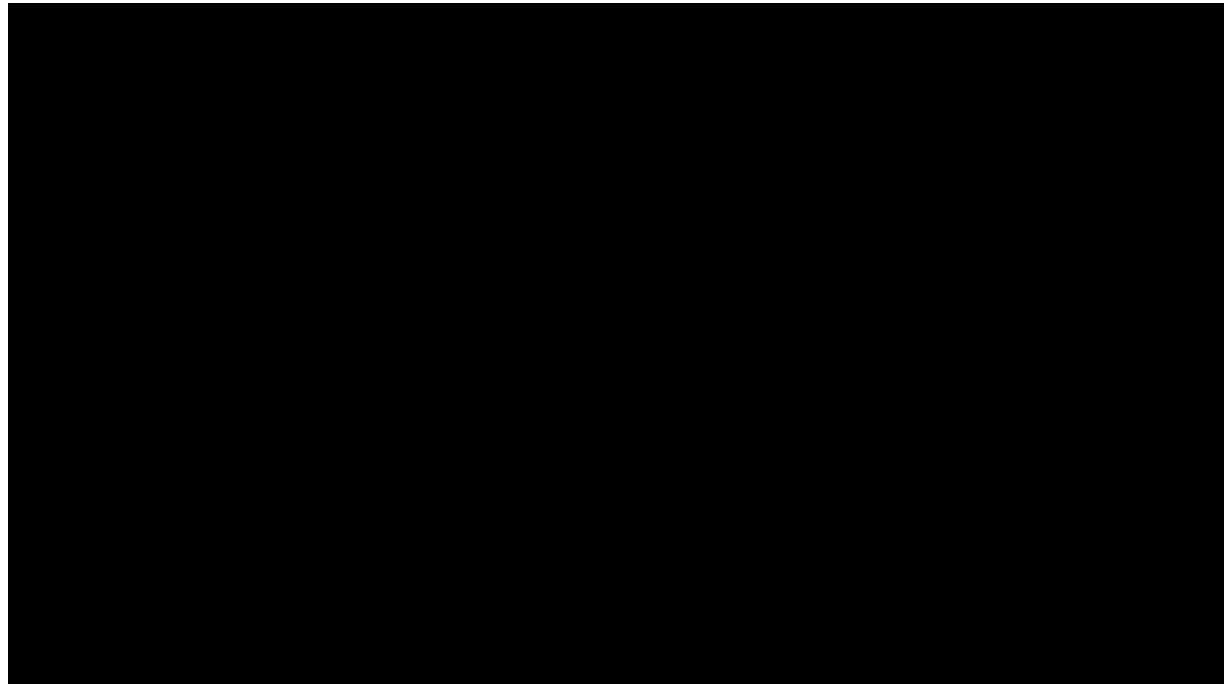
VISUALISIERUNG
VON MATERIAL

Quelle: <https://romereborn.org/content/roman-forum>



Minecraft Palmyra

Nachbildung der antiken Stadt Palmyra in Minecraft (Universität Leiden)



VISUALISIERUNG
VON MATERIAL

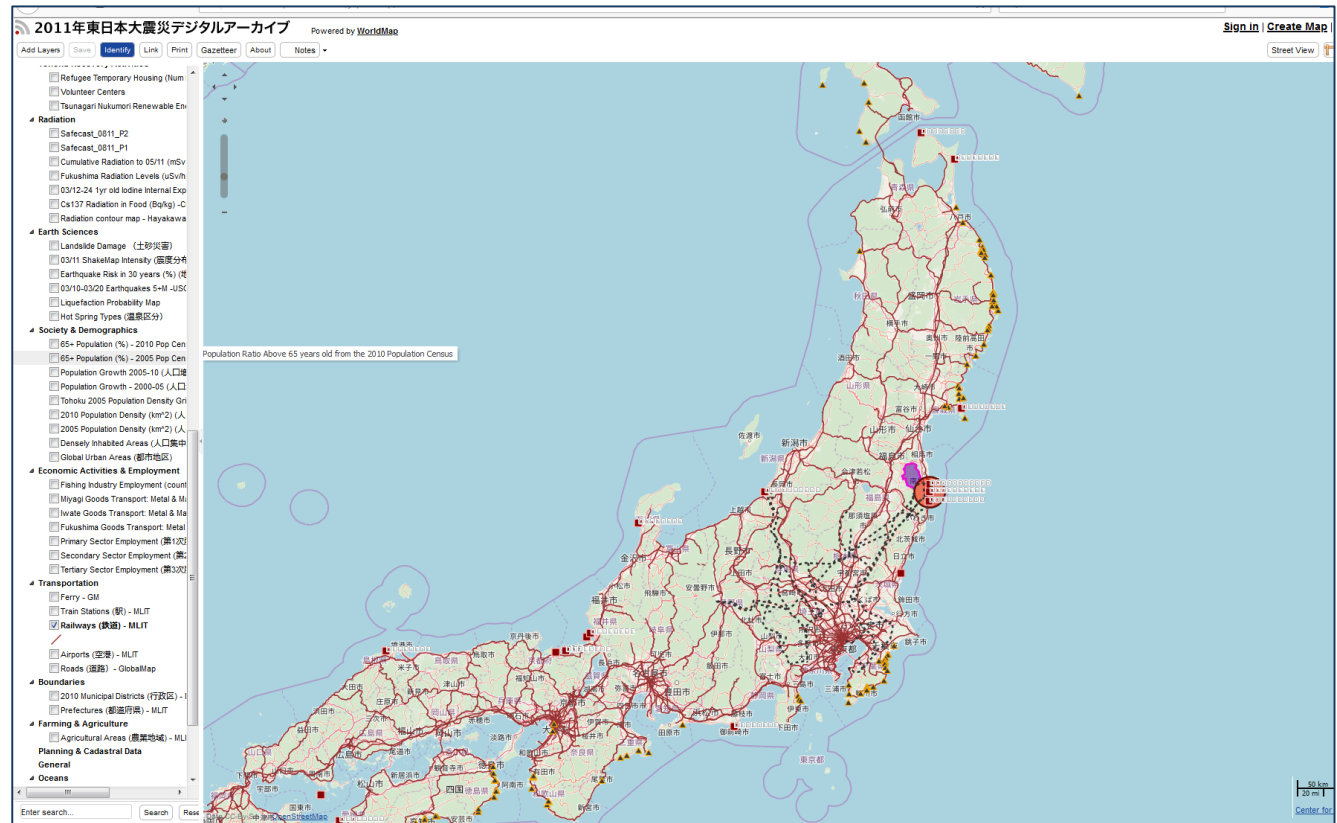
Quelle: <https://www.universiteitleiden.nl/en/news/2019/11/rebuilding-palmyra-in-minecraft>

GEOGRAPHICAL INFORMATION SYSTEMS (GIS)



Harvard World Maps

GIS-Projekt der Universität Harvard zur Visualisierung von geographisch abbildbarer Daten bei gleichzeitiger Dokumentation der Daten



<http://worldmap.harvard.edu/japanmap/>

VISUALISIERUNG
VON DATEN



HARVARD
 Dataverse

Search ▾ About User Guide Support Sign Up Log In

Skinner Regional Systems Analysis Dataverse (Regional Systems Analysis Project) Home Page

Harvard Dataverse > **Skinner Regional Systems Analysis Dataverse**

[Contact](#) [Share](#)

An archive of the datasets developed for the **Regional Systems Analysis** laboratory, under the direction of G. William Skinner. See also [sample layers as webmaps](#).

The core datasets, including GIS layers and tabular data, are being preserved in this repository. Together the datasets will present a case study of the methodology used for determination of the Levels in Urban Hierarchy (LUH), Urban-Rural Continuum values (URC), and the delineation of Core Periphery Zones (CPZ).

For a comprehensive description of the analysis method, please see *Spatial Contexts of Urbanization and Land Conservation in North China* by Mark Henderson.

Archive Editors:
[Mark Henderson](#), Mills College Public Policy Program
[Merrick Lex Berman](#), China Biographical Database

China - Modern Dataverse

China - Late Imperial Dataverse

Japan Dataverse

France Dataverse

[Advanced Search](#)

☒ **Dataverses (4)**
☒ **Datasets (26)**
☐ **Files (718)**

Dataverse Category
[Research Project \(4\)](#)

Publication Year
 2012 (6)
 2013 (6)
 2018 (6)
 2016 (5)
 2017 (3)
[More...](#)

Subject
[Social Sciences \(10\)](#)
[Earth and Environmental Sciences \(4\)](#)

1 to 10 of 30 Results

China-CPZ (Core Periphery Zones)
 May 10, 2018 - China - Modern Dataverse
 G. W. Skinner; Zumou Yue; Mark Henderson, 2013, "China-CPZ (Core Periphery Zones)", <https://doi.org/10.7910/DVN/HBERON>, Harvard Dataverse, V2
 China-CPZ, the core periphery zones of China, circa 1990. Core-Periphery Zones are based on a regional systems analysis of China, drawing on data about population density and agricultural intensity for county level units (see the Regional Systems Index in the ChinaA datafile). Zo...

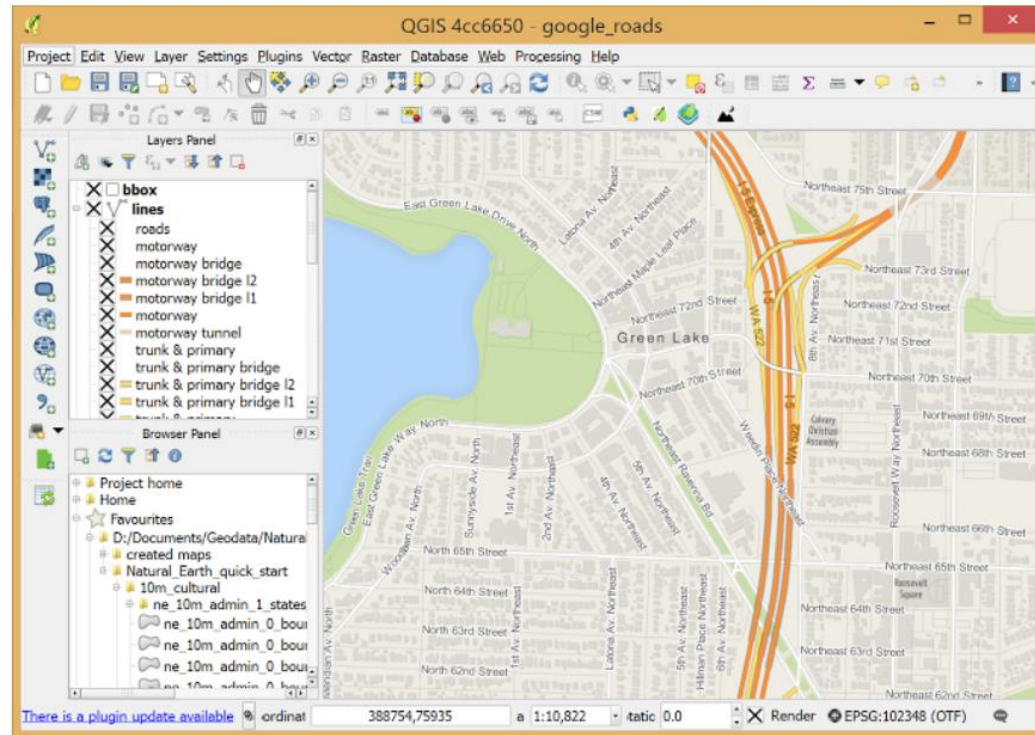
Key to Skinner Dataset Names and Abbreviations
 May 10, 2018
 Henderson, Mark; Skinner, G. W., 2018, "Key to Skinner Dataset Names and Abbreviations", <https://doi.org/10.7910/DVN/T7JSDA>, Harvard Dataverse, V1
 Abbreviations for datasets with descriptions of the contents, time periods, and areas covered by the Skinner datasets.

Dataset (1990) Central Place Analysis
 May 10, 2018 - China - Modern Dataverse



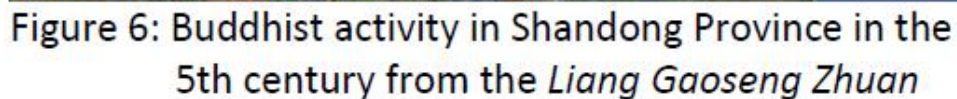
QGIS

Open-Source geographisches Informationssystem zum Mappen von Daten in geographische Dimension



QGIS ist ein benutzerfreundliches Open-Source geographisches Informationssystem (GIS), das unter der GNU General-Public-License steht. QGIS ist ein offizielles Mitglied der

VISUALISIERUNG
VON DATEN



1. To provide researchers with tools that can answer questions which cannot be asked of traditional reference tools, and to visualize these answers in a GIS-like manner. These might be query parameters such as: Show on a map who was active in Zhejiang province during the Tang dynasty (618-907), according to collection Y; or: Show all places the monk Xuanzang passed through on his way to India, according to collection Z.
2. To create an interface where researchers and the interested public can read the biographies and have immediate access to information about persons, places and dates. This is achieved by embedding Google Earth to create interactive maps and connecting person and place names to an authority database.
3. To provide an open-licensed corpus, markup from which can be used and modified by other researchers to interrogate the texts in new ways.

*TEI = *Text Encoding Initiative*

VISUALIZATION OF DATA

NETWORK-ANALYSIS



Investigating the Relationships between Scholars and Politicians in Ancient China: Taking the Yuanyou Era as an Example

Shang Wenyi* and Huang Winbin*

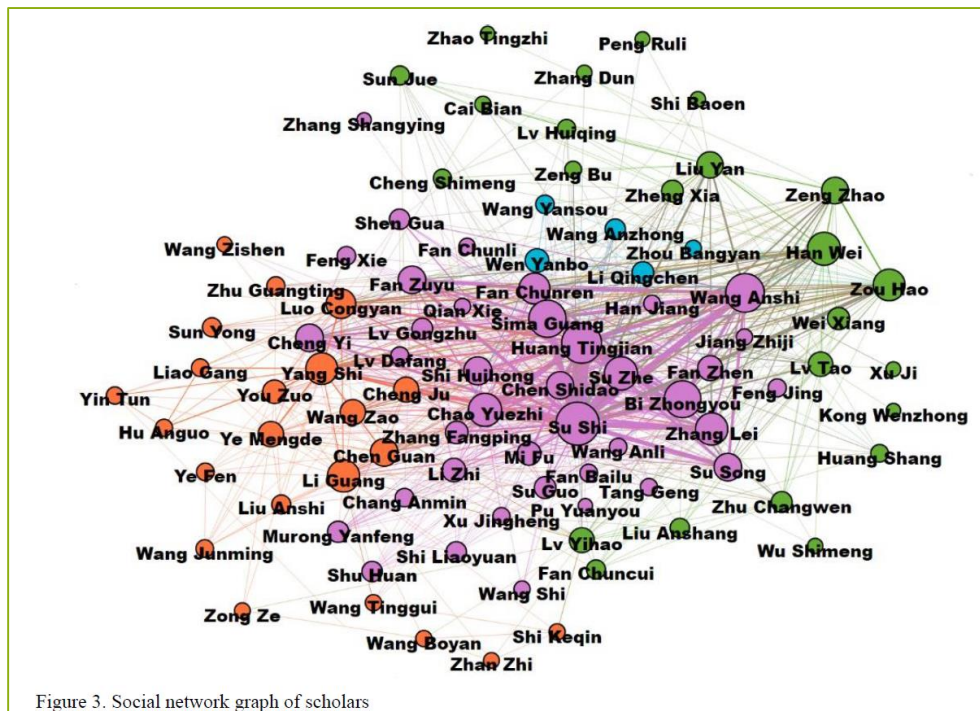
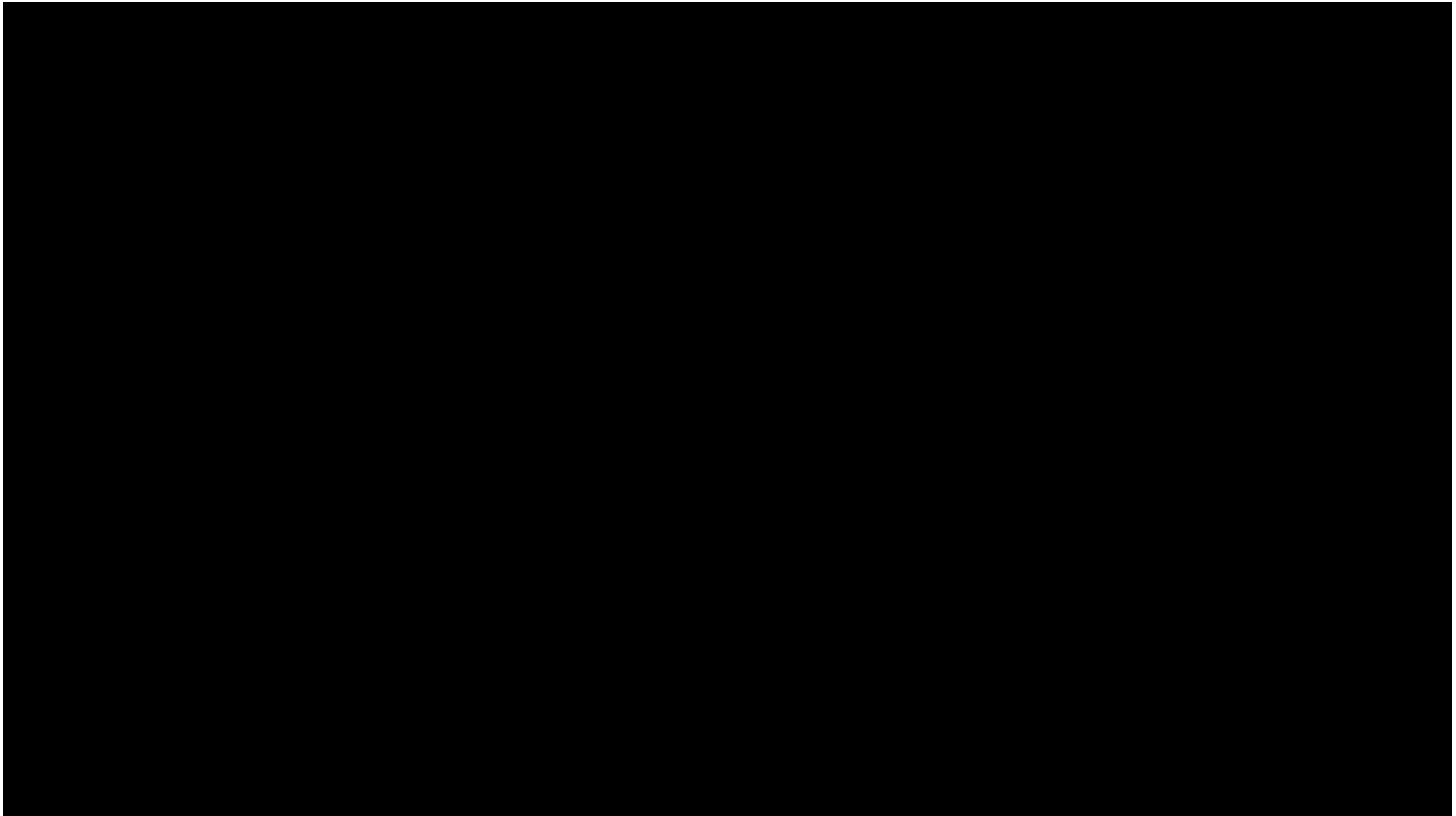


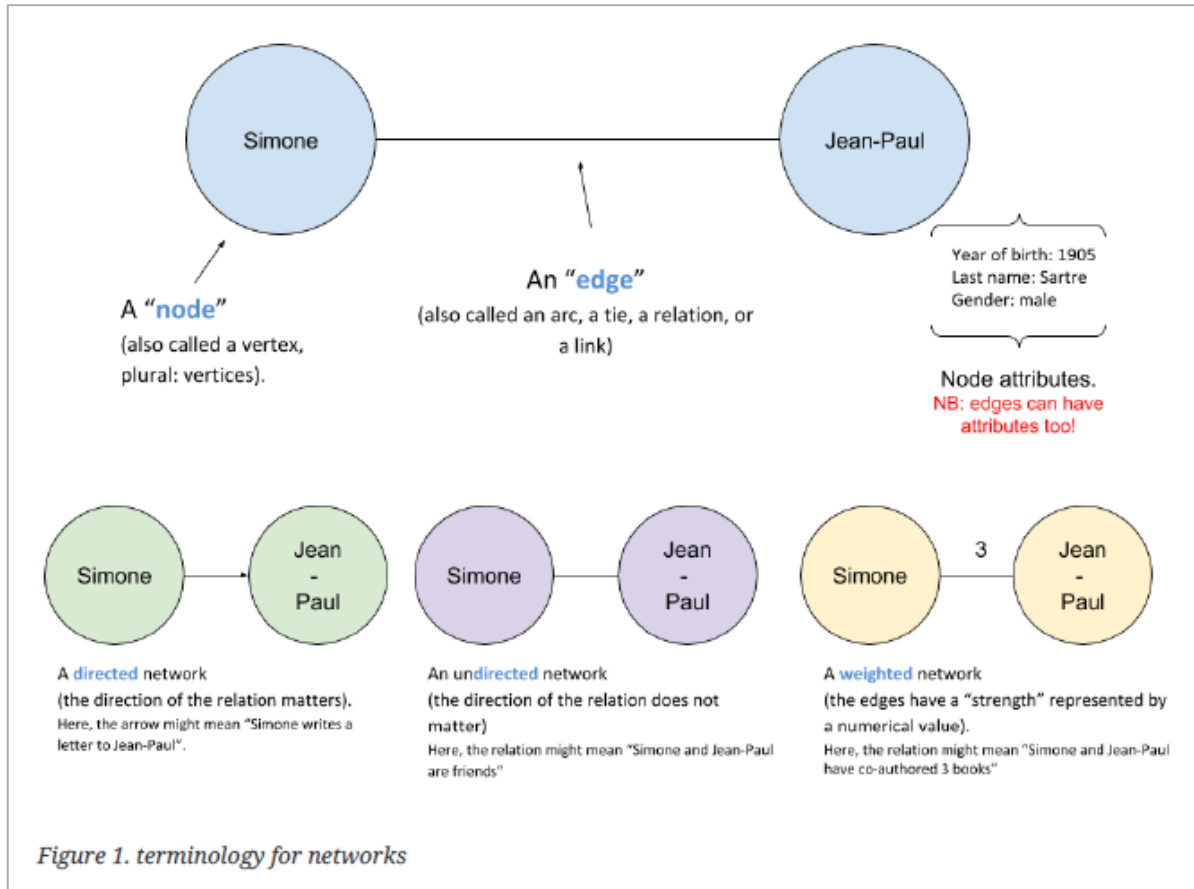
Figure 3. Social network graph of scholars

Conversational Network in the Chinese Buddhist Canon

[illegible]



<https://www.youtube.com/watch?v=gcfAT8aMxuQ&t=259s>



- Degree
- Density
- Centrality
- ✓ *Closeness centrality*
- ✓ *Degree centrality*
- ✓ *Betweenness centrality*

Quelle: „Simple Gephi Project from A to Z“ (2017)

Infos zu Centrality measures: <https://www.youtube.com/watch?v=89mxOdWPfxA>

JETZT SIND SIE DRAN!

Verschiedene Visualisierungsprojekte kennenlernen



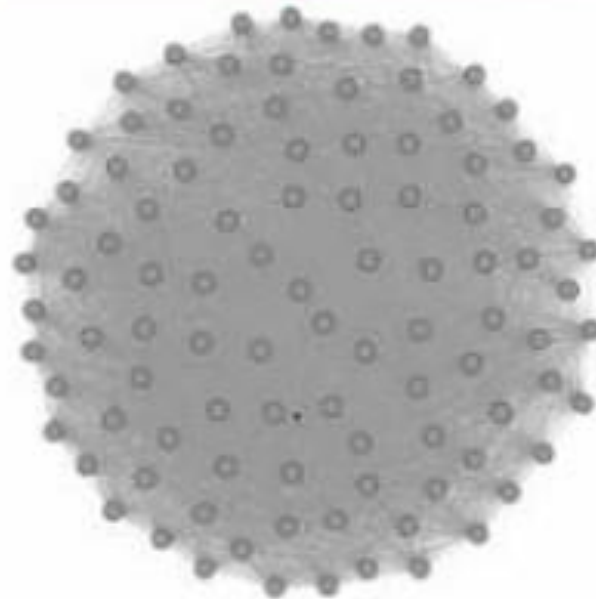
- BEISPIEL 1 | GEPHI UND NETZWERKANALYSE

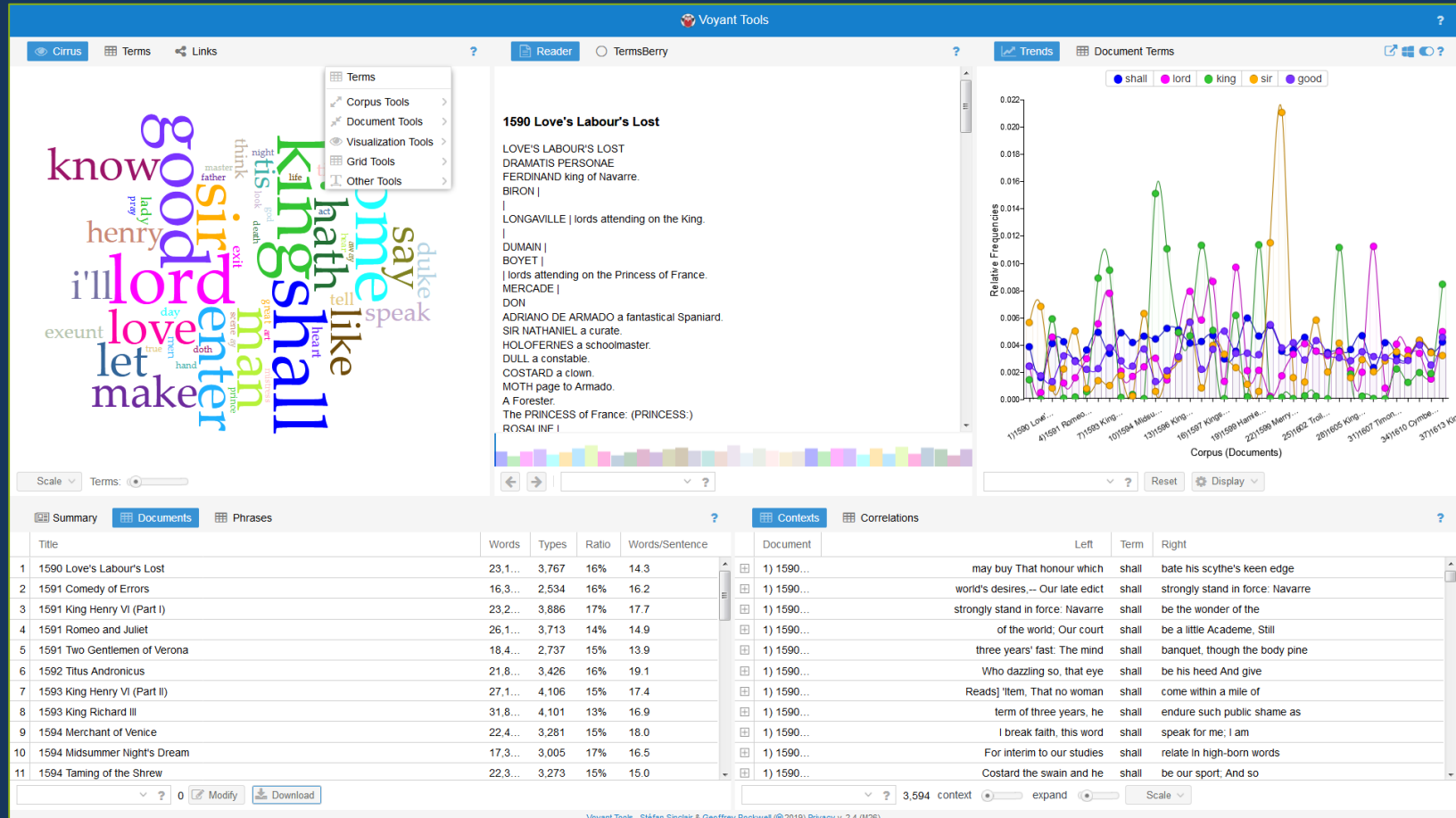
- BEISPIEL 2 | VOSVIEWER UND BIBLIOMETRISCHE NETZWERKANALYSE

- BEISPIEL 3 | CTEXT-TEXTTOOLS, TEXT-REUSE & NETZWERKANALYSE



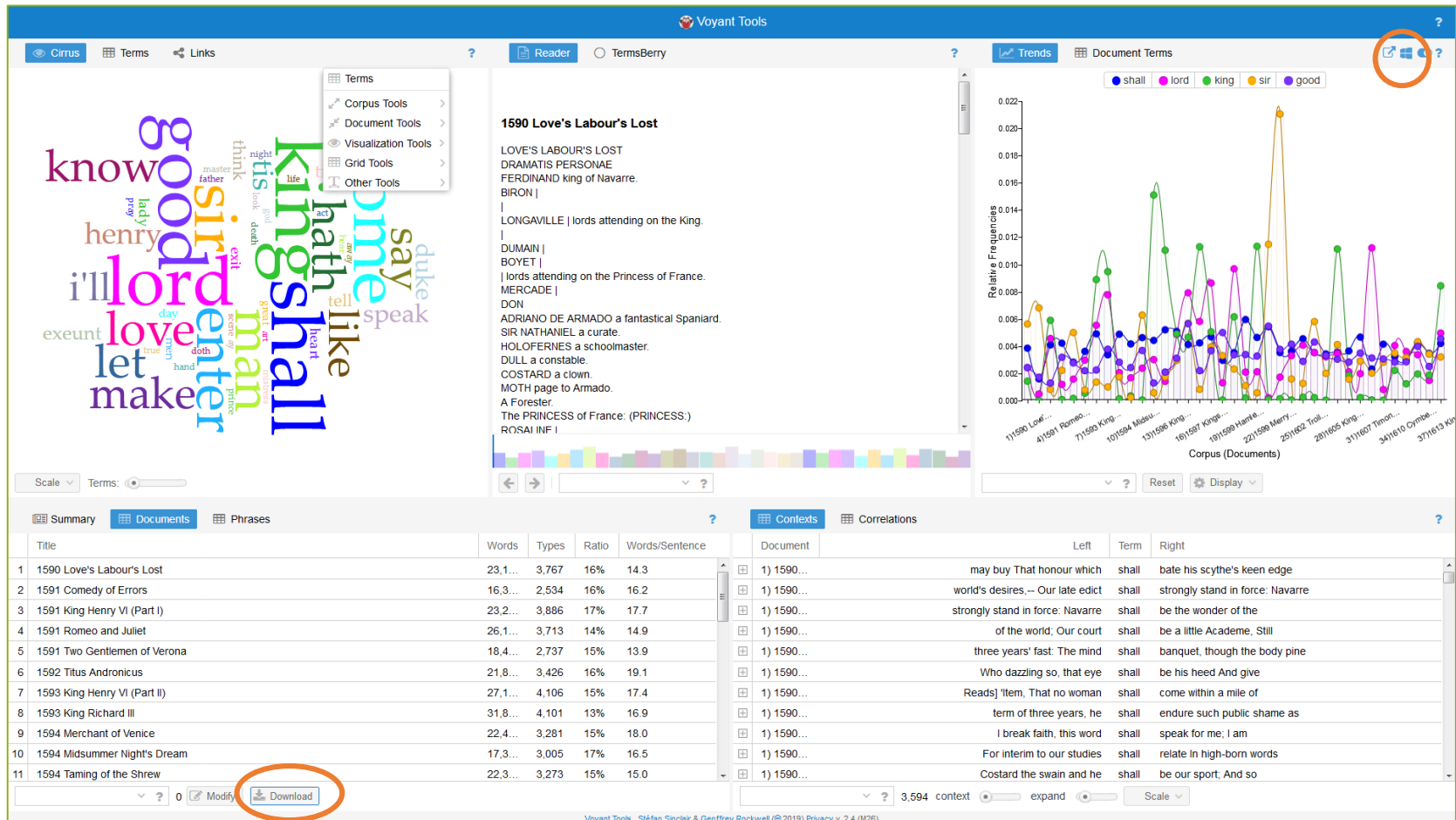
Example: Senate Voting Records

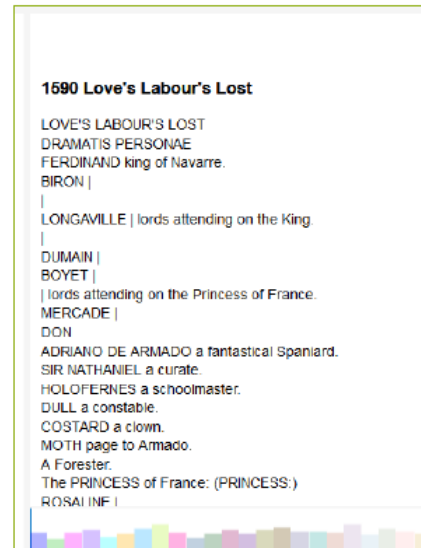
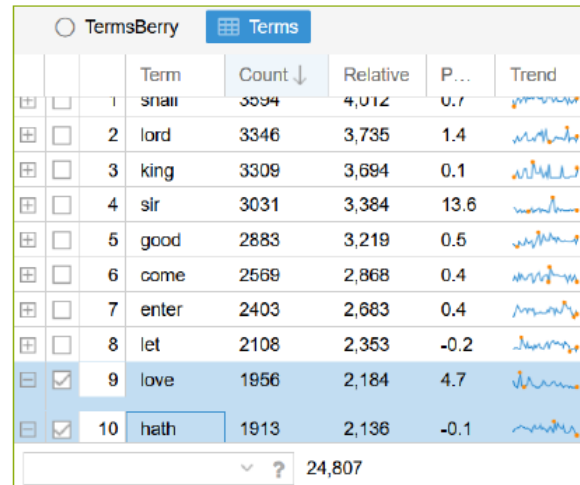






- Metadaten downloaden, URLs exportieren, Tools auswählen, Scale festlegen, Einstellungen anpassen, ...

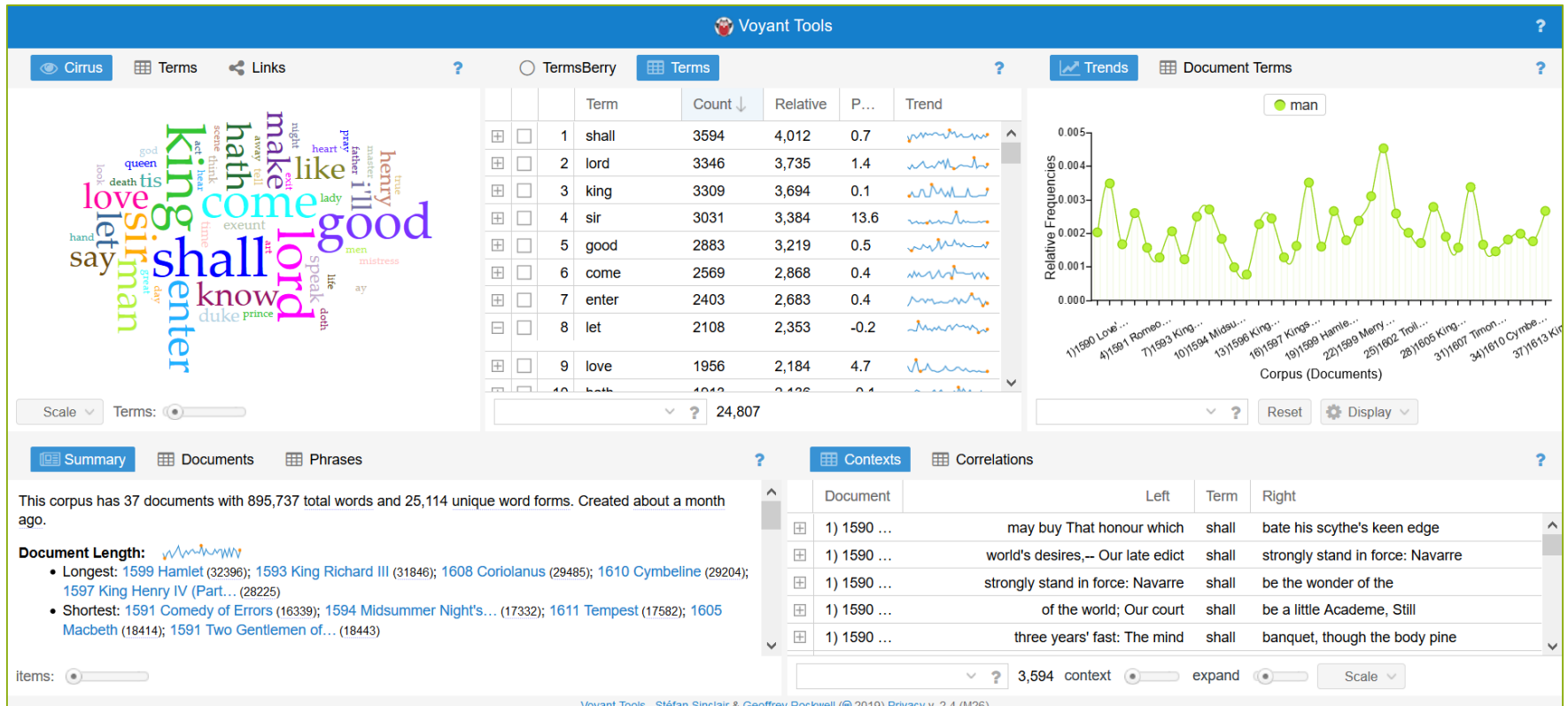




- Reader
- Summary
- Cirrus
- Terms
- Trends
- Bubblelines
- Document Terms
- Contexts
- Collocates
- Correlations (?)
- Phrases



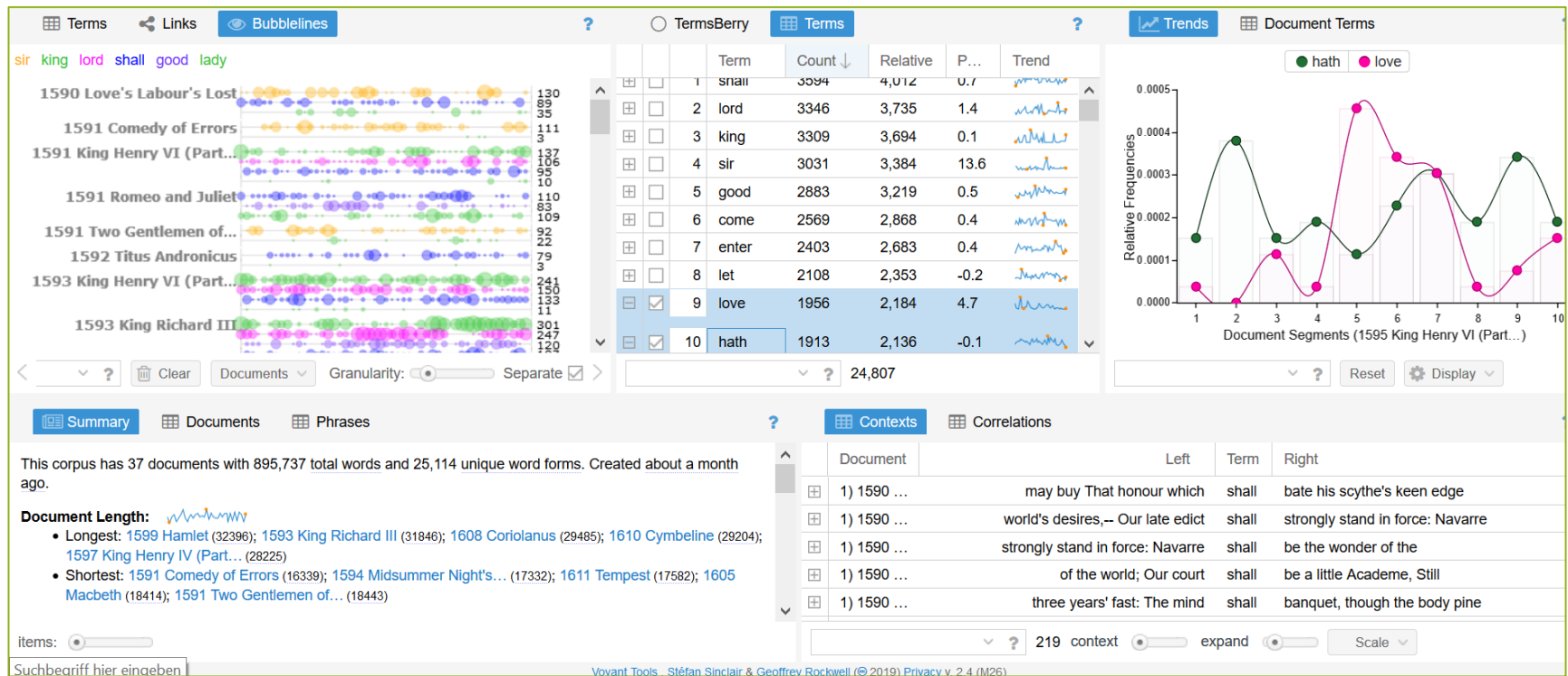
CORPUS TOOLS → TERMS & TRENDS



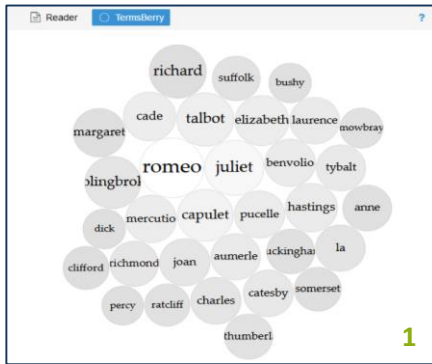
Quelle: <https://voyant-tools.org/?panels=cirrus%2Ccorpus%2Cterms%2Ctrends%2Csummary%2Ccontexts&corpus=shakespeare>



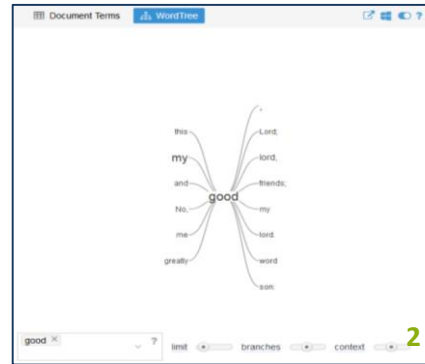
CORPUS TOOLS → BUBBLE LINES



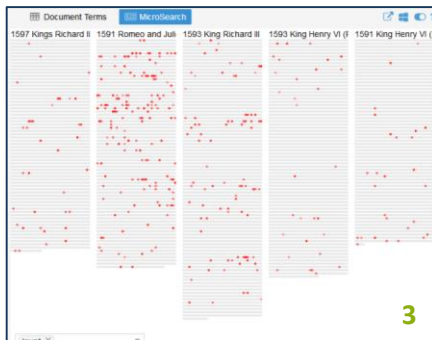
Quelle: <https://voyant-tools.org/?panels=bubblelines%2Ccorpus%2Cterms%2Ctrends%2Csummary%2Ccontexts&corpus=shakespeare>



1



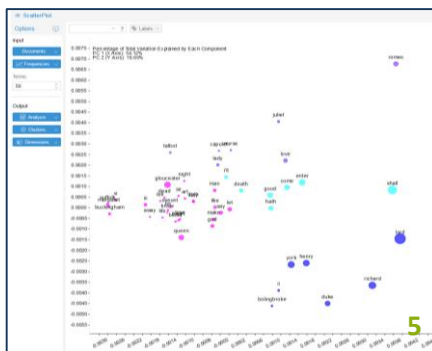
2



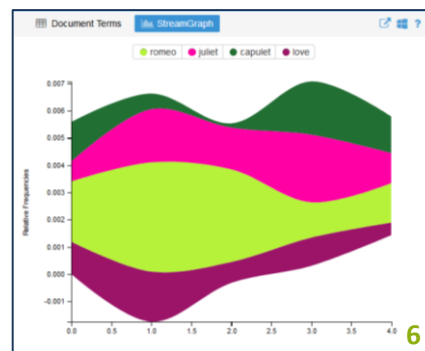
3



4



5



6

1. [TermsBerry](#)
2. [WordTree](#)
3. [MicroSearch](#)
4. [Mandala](#)
5. [ScatterPlot](#)
6. [StreamGraph](#)

<https://voyant-tools.org/?corpus=0cec7de2152125145ff7272c08879e41&view=scatterplot&panels=correlations,reader,trends,documents,contexts>



- **METADATEN -> Summary**

- ✓ Items
- ✓ Frequency
- ✓ Type / Token ratio = Vocabulary density (*number of words in the document to the number of unique words in the document*)
- ✓ *Distinctive words (= keywords)*

- **WICHTIGE BEGRIFFLICHKEITEN**

- ✓ **Stopwords** (In Voyant Tools you can select from pre-existing stopwords lists in various languages or create your own. For tools that support stopwords lists, you can begin by clicking on the options icon (the may not appear if it's a tool that doesn't support stopwords) (<https://voyant-tools.org/docs/#!/guide/stopwords>))
- ✓ **Relative Frequency**
- ✓ **Peakness**
- ✓ **TF-IDF** (Text frequency-inverse document frequency)



■ SUCHABFRAGEN-BEISPIELE

- ✓ **love** match exact term love
- ✓ **love*** match terms that start with the prefix love and then a wildcard as one term
- ✓ **^love*** match terms that start with love as separate terms (love, lovely, etc.)
- ✓ ***ove** match terms that end with the suffix ove as one term
- ✓ **^*ove** match terms that end with suffix ove as separate terms (love, above, etc.)
- ✓ **love,hate** match each term separated by commas as separate terms
- ✓ **love|hate** match terms separated by pipes as a single term
- ✓ **"love him"** love him as an exact phrase (word order matters)
- ✓ **"love him"~0** love him or him love phrase (word order doesn't matter but 0 words in between)
- ✓ **"love her"~5** match love near her (within 5 words)
- ✓ **^love*,love|hate,"love her"~5** combine syntaxes

Quelle: <https://voyant-tools.org/docs/#!/guide/search>

INTERNETQUELLEN

- **Gephi:** <https://gephi.org/>
- **VOS Viewer:** <http://www.vosviewer.com/>
- **QGIS:** <https://qgis.org/de/site/Open Source>
- **GRASS GIS** (Free): <https://grass.osgeo.org/screenshots/user-interface/>
- **ArcGIS:** <https://www.arcgis.com/index.html>
- **MAXQDA:** <https://www.maxqda.de/funktionsumfang>
- **Spatial History Project** (Stanford): <https://web.stanford.edu/group/spatialhistory/cgi-bin/site/projects.php>
- **Literaturkarte Ruhr:** http://www.literaturkarte.ruhr/map_f_all.php
- **China Historical GIS Project** (Harvard): <http://chgis.fas.harvard.edu/pages/intro/>
- **Japan / China WorldMaps:** <https://worldmap.harvard.edu/maps/china-history>
- **Wortverläufe im DWDS** (deutsch):
<https://www.dwds.de/r/plot?view=2&corpus=zeitungen&norm=abs&smooth=line&genres=0&grand=1&slice=1&prune=0&window=0&wbase=0&logavg=0&logscale=0&xrange=1945:2017&q1=China>
- **Kollokationen in DiaCollo** (deutsch):
<http://kaskade.dwds.de/dstar/zeit/diacollo/?query=Umwelt&date=&slice=10&score=ld&kbest=10&cutoff=&profile=2&format=cloud&groupby=&eps=0>
- **CBETA** (Chin., buddhistische Texte): http://cbetaonline.dila.edu.tw/zh/T0001_001
- **BCC** (Chin., Zeitungstexte): <http://bcc.blcu.edu.cn/lang/zh>
- **GoogleMyMap „Pilgrimage of Xuanzang“:**
https://www.google.com/maps/d/viewer?mid=1aA0N5NQAgba2cKeY0_Z4Kk3pSsM&ll=32.18462061496944%2C79.5042519162231&z=4
- **Sturgeon, Textual Reuse:** <https://youtu.be/uLTP6pO2MtU>
- **Chinese Text Project:** ctext.org
- **Visual Linguistics Project:** https://www.bubenhofer.com/graphics/bubenhofer_visual_linguistik.mp4
- **GeoCollocations:** <http://www.bubenhofer.com/geocollocations/>
- **RUB Zentrum für Geoinformation:** <http://www.izg.rub.de/>