

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*

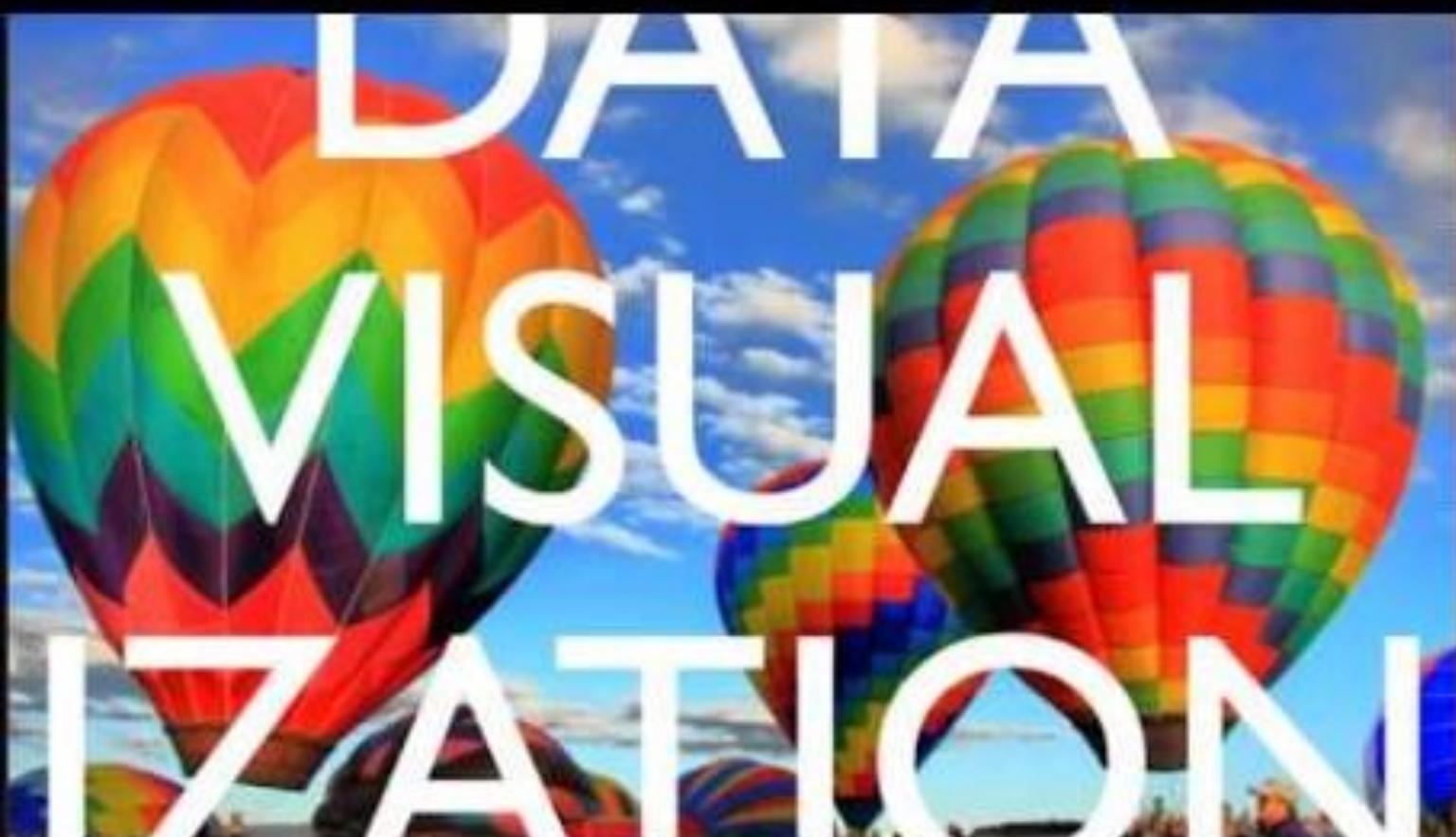
# VISUALISIERUNG GEGENSTAND DER VISUALISIERUNG



VISUALISIERUNG  
VON TEXTEN

VISUALISIERUNG  
VON MATERIAL

VISUALISIERUNG  
VON DATEN



Quelle: Jengolbeck, <https://www.youtube.com/watch?v=tOB2sxlumh0>



▪ **Information visualization**

▪ **Scientific visualization**

▪ **Infographics**

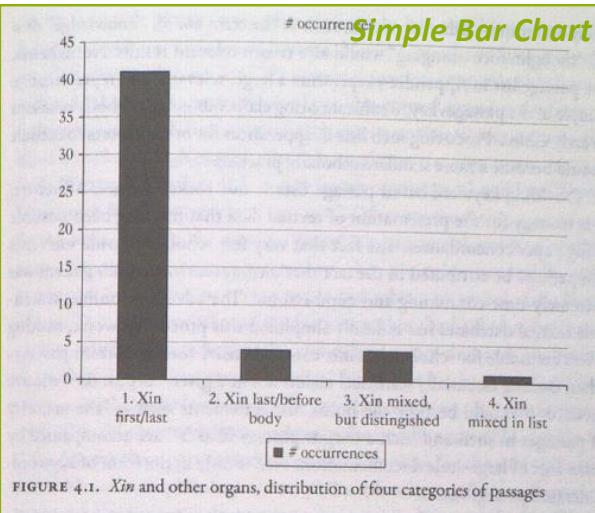
▪ **Static visualization**

▪ **Dynamic visualization**

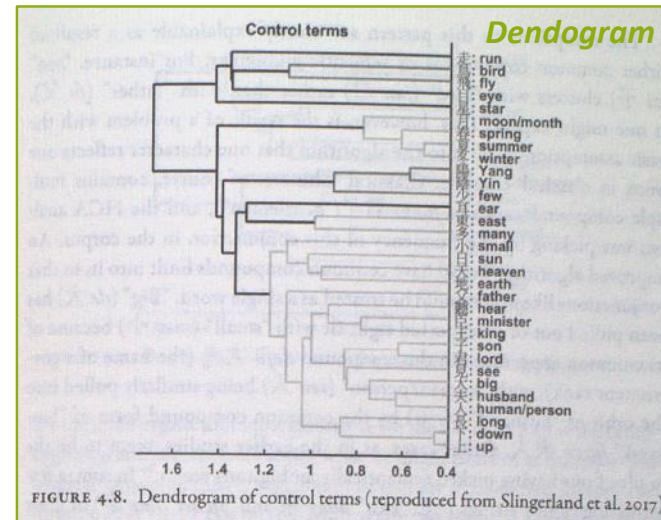
▪ **Interactive visualization**



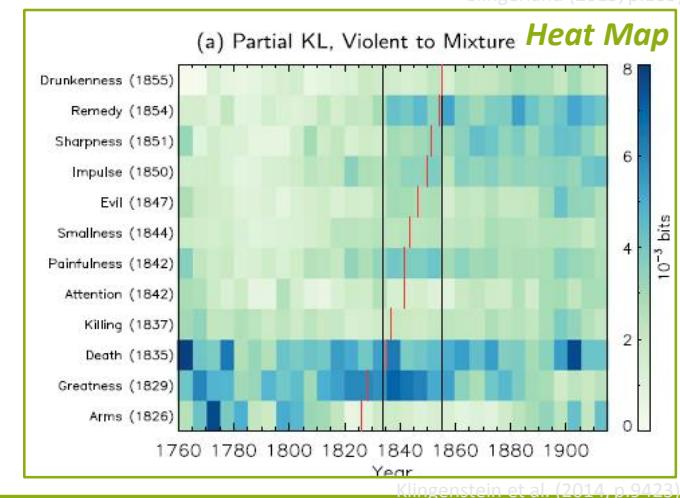
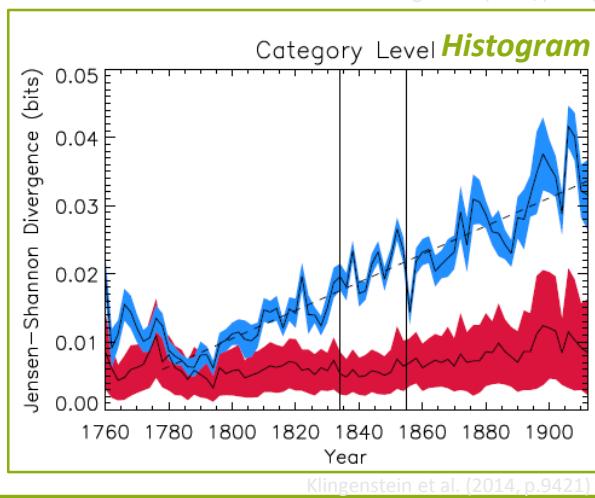
## Static visualizations

VISUALISIERUNG  
VON DATEN

Slingerland (2019, p.149)



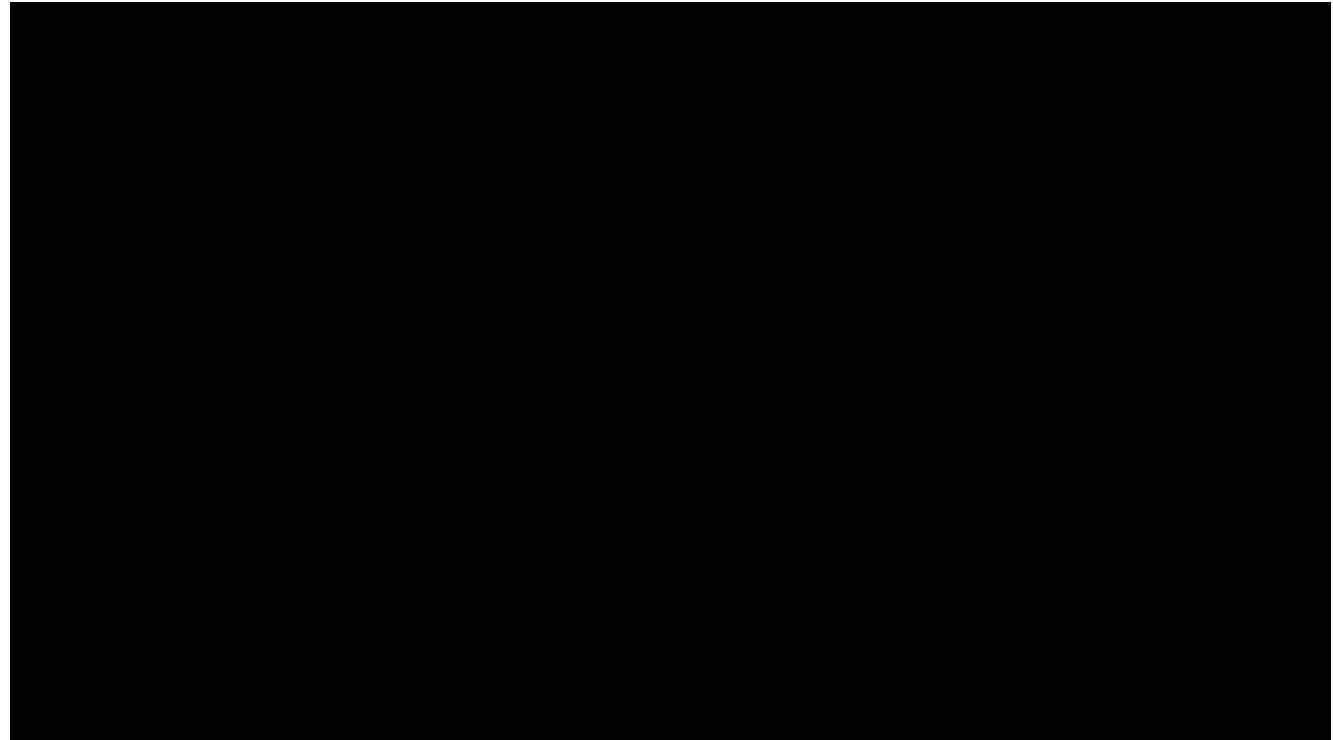
Slingerland (2019, p.185)



## VISUALISIERUNG VON DATEN

*Dynamic visualizations*

# VISUALISIERUNG TYPEN DER VISUALISIERUNG | BEISPIELE



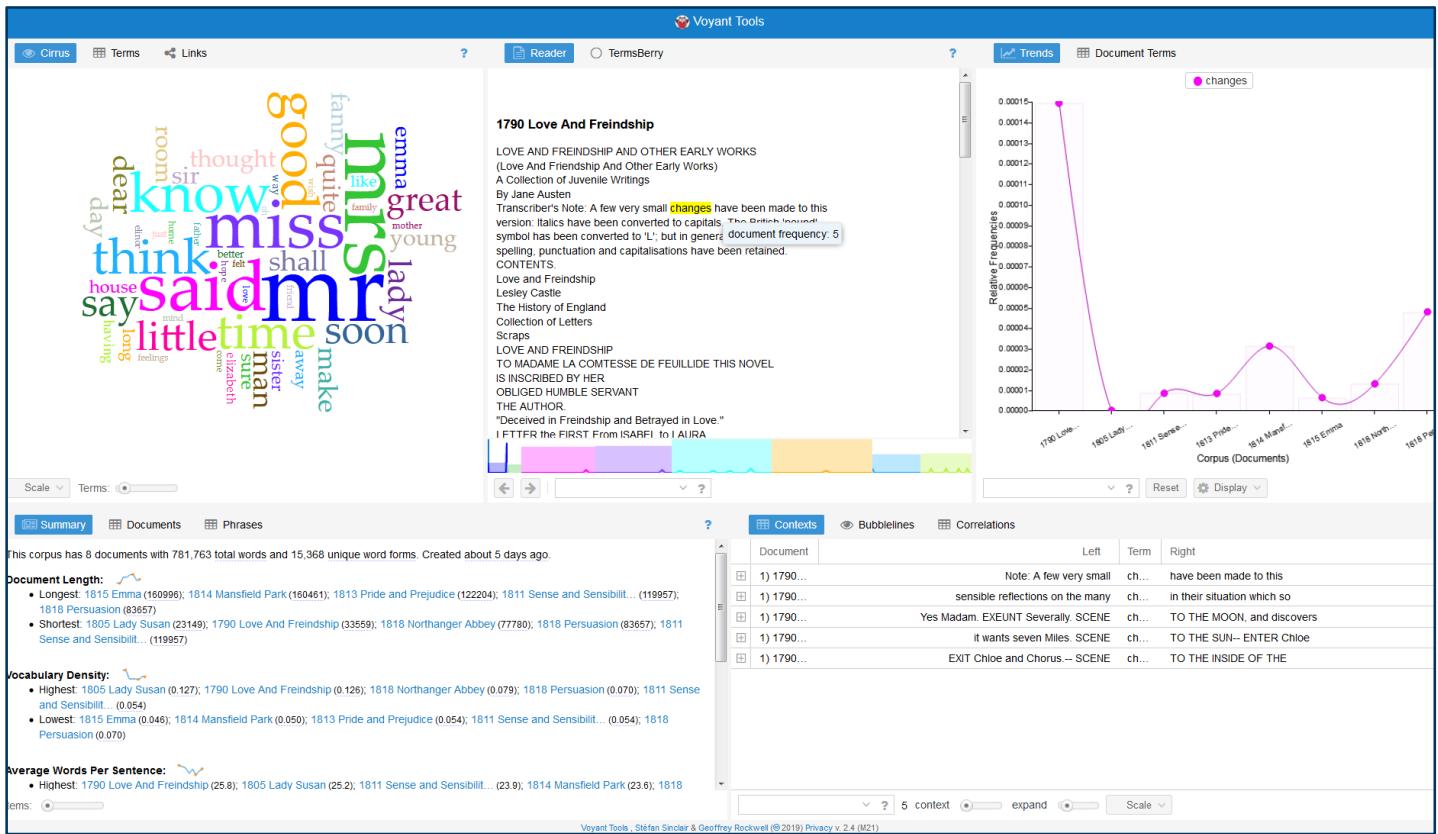
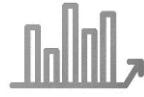
China Biographical Database Project: <https://projects.iq.harvard.edu/cbdb/home>

## Interactive visualizations

VISUALISIERUNG  
VON DATEN

# VISUALISIERUNG

## TYPEN DER VISUALISIERUNG | BEISPIELE



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

COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN.

RUHR  
UNIVERSITÄT  
BOCHUM

RUB

# VISUALIZATION / DIGITIZATION OF TEXTS

# VISUALISIERUNG GEGENSTAND DER VISUALISIERUNG



## Ten Thousand Rooms Project

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

The screenshot shows a digital interface for visualizing classical Chinese texts. On the left, a sidebar lists chapters of 'Kongzi Jiayu' (孔子集語). The main area displays a page from the book with vertical columns of Chinese text. A blue box highlights a specific section of text. To the right, a panel titled 'Annotations' shows a list of available annotations: Transcription, Analytical Commentary, Textual Commentary, Translation, Project Log, and a link to 'Viewed Documents, "Overview & Details"'. A detailed translation of the highlighted text is provided in English:

1.2 Elder Dai's Ritual Records, "An Exhortation to Learn" [chapter 64]:  
Kongzi said: Li [i.e., Kongzi's son Boyu]: It is not permissible for a noble man not to learn, nor is it permissible for a noble man to learn, nor is it permissible for him not to adorn himself when appearing before others. Not adorning himself he would lack the proper appearance, lacking the proper appearance he would not be respectful, not being respectful he would not establish himself. "Adornment" is possessing brilliance from afar; "learning" is enhancing one's brilliance close at hand. It is like a polluted body of water when rainwater gathers and the cattails begin to grow. Looking at it from above, who can tell that it is not really a spring?

Transcription:  
ir a noble man not to learn, nor is it itself when appearing before others. Not adorning himself he would lack the proper appearance, lacking the proper appearance he would not be respectful, not being respectful he would lack ritual propriety, and lacking ritual propriety he would not establish himself. "Adornment" is possessing brilliance from afar; "learning" is enhancing one's brilliance close at hand. It is like a polluted body of water when rainwater gathers and the cattails begin to grow. Looking at it from above, who can tell that it is not really a spring?

Analytical Commentary:  
Transliteration:  
Translation:  
Project Log:  
Viewed Documents, "Overview & Details":

VISUALISIERUNG  
VON TEXTEN

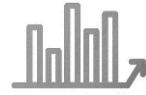
Quelle: <https://tenthousandrooms.yale.edu/node/311/mirador?canvas=14116>

COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN.

RUHR  
UNIVERSITÄT  
BOCHUM

RUB

# VISUALISIERUNG GEGENSTAND DER VISUALISIERUNG



## SAT

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

The screenshot shows the SAT digital library interface. At the top, there's a search bar and navigation links for 'AT大藏經DB' (Collection Database), '目次' (Table of Contents), '人名' (Name), '検索オプション' (Search Options), '1884', '検索' (Search), and '回像検索' (Image Search). Below the search bar, it says '浄土全書DB: 全0件' (Pure Land Complete Collection DB: 0 items). The main area displays a digital page from the Chinese Buddhist Canon (No. 1884) titled '註華嚴法界觀門'. The page contains dense Chinese text. To the right of the page, there's a sidebar with '履歴' (History), 'テキスト分析' (Text Analysis), '華嚴五教止觀' (Hua Yan Five Teachings), '起信論海東疏刊行序' (Preface of the Edition of the Sea of the Kegon), and '起信論海東疏刊行序' (Preface of the Edition of the Sea of the Kegon). Below the sidebar, there's a detailed description of the text, including '注華嚴法界觀門序' (Preface of the Hua Yan Five Teachings), '唐綿州刺史裴休述' (Written by Cui Xiu, Governor of Jinzhou), and a long quote in Chinese. At the bottom of the page, there are small thumbnail images of other pages from the collection.

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

## VISUALISIERUNG VON TEXTEN

COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN.

RUHR  
UNIVERSITÄT  
BOCHUM

RUB

# VISUALISIERUNG GEGENSTAND DER VISUALISIERUNG



## CBETA

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

**VISUALISIERUNG  
VON DATEN**

**VISUALISIERUNG  
VON TEXTEN**

**T45n1867 華嚴五教止觀** / 卷/篇章一

捜索 : 治

#

1 而裁方。病盡而方息。**治**執而施藥。執造而藥已  
2 破此病故。都闢四藥以**治**四病。其中別門各有藥  
3 六界。謂眼界等是也。**治**前計我之病也。二分外  
4 六界。謂色界等是也。**治**前計我所之病。三分中  
5 譏界耳識界等者是也。**治**前我見聞等病。是已上  
6 用此為一百八煩惱。所**治**之病既癥。能  
7 之病既癥。能**治**之藥亦然。俱根塵識等  
8 觀也。問作如是觀者。**治**等病法耶。答  
9 何等病耶。答**治**上執法之病。何者。法

卷/篇章 : [1]

No. 1867

[3]五教止觀一乘十玄門合行敘

夫五教止觀也者。華嚴初祖帝心尊者所造。且一乘十玄門也者。第二祖至祖大師所撰也。傳釋于大不思議經。而始起五敘。迺立十玄。是顧於無間重圓因陀羅珠網之寶冊。示乎圓融隱銳光頗黎鏡之玉章也。或曰書林井氏袖彼二小策子來謂云。頂曰影刻入梓。因思欲及于遐邇。而卷軸至少。若別行者恐損披閱。是故合集以行焉。因乞措一解於卷首。茲第應索書而為敘。

惟時元祿第九丙子八月毅旦審智積覺眼空敬

華嚴五教止觀

京終南山文殊化身 杜順 說

行人修道簡邪入正止觀法門有五。

一法有我無門(小乘教) 二生即無生門(大乘始教) 三事理圓融門(大乘終教) 四語報雙絶門(大乘頓教) 五華嚴三昧門(一乘圓教)

第一法有我無門

夫對病而裁方。病盡而方息。**治**執而施藥。執造而藥已。為病既多。與藥非一。隨機進修異。所以方便不同。今偏就五停心中。為眾生著我者。說界分別觀。眾生從無始已來。執身為一。計我我所。然計我有二種。一即身執我。二離身執我。言離身執我者。謂外道計身內別有神我者是也。廣如經論中破。於此不更繁文。言即身執我者。**[4]**執我如來慈悲為破此病故。都開四藥以**治**四病。其中別門各有藥病。具如後釋。言四病者。一執身為一我。二執四大。三執五陰。四執十二入。言四藥者。一色心兩法。二四大五陰。三十二入。四十八界是也。次釋。若眾生執身為一而我成病者。即說色心二法為藥。亦云。此中乃有色心二法。云何為一我耶。眾生聞此遂即轉執色心為實成病。即為開一色。**[5]**即為開一心。色為四色。即四大是也。開一心為四心。即五陰中四陰是也。此乃是四色四心。云何但執一色一心為一我耶。眾生又即轉執四色四心成病。佛即為合四大為一色。即五陰中色陰是也。合四心為一心。即十二入中意入是也。眾生聞此又更轉執成病。佛即為分一色為

目錄 卷/篇章 搜尋

詞頻統計資訊

部類	身
同合部類	10393
本緣部類	10693
般若部類	38472
法華部類	54510
華嚴部類	54437
寶積部類	10868
涅槃部類	13931
大集部類	5808
經集部類	57236
密教部類	49578
律部類	36882
毘曇部類	31613
中觀部類	3275
瑜伽部類	32676
論集部類	9191
淨土宗部類	19050
禪宗部類	64294
史傳部類	31697

詞頻在各朝代分布

朝代	身
東漢	2219
曹魏	121
吳	2191
西晉	6542
晉	224
前涼	19
東晉	9037
前秦/苻秦	5129
後秦/姚秦	16533
西秦/乞伏秦	319
北涼	8500
劉宋	8131
北魏/元魏/後魏	5608
蕭齊/南齊	601
南北朝	513
梁/蕭梁	6151
東魏	4489
西魏	0

Quelle: [http://cbetaonline.dila.edu.tw/zh/T1867\\_001](http://cbetaonline.dila.edu.tw/zh/T1867_001)

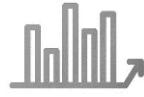
COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN.

RUHR  
UNIVERSITÄT  
BOCHUM

RUB

# VISUALIZATION / DIGITIZATION OF MATERIAL

# VISUALISIERUNG GEGENSTAND DER VISUALISIERUNG



## Xiangtangshan Caves

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

**响堂山石窟项目**  
**Xiangtangshan**  
caves project

Home    Caves »    Sculptures    3D Models »    Exhibitions    Technology    Institutions    People    Support

**CAVES**

The Chinese Buddhist cave temples of Xiangtangshan, "Mountain of Echoing Halls," from the Northern Qi dynasty (550-577)

VISUALISIERUNG  
VON MATERIAL

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

COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN.

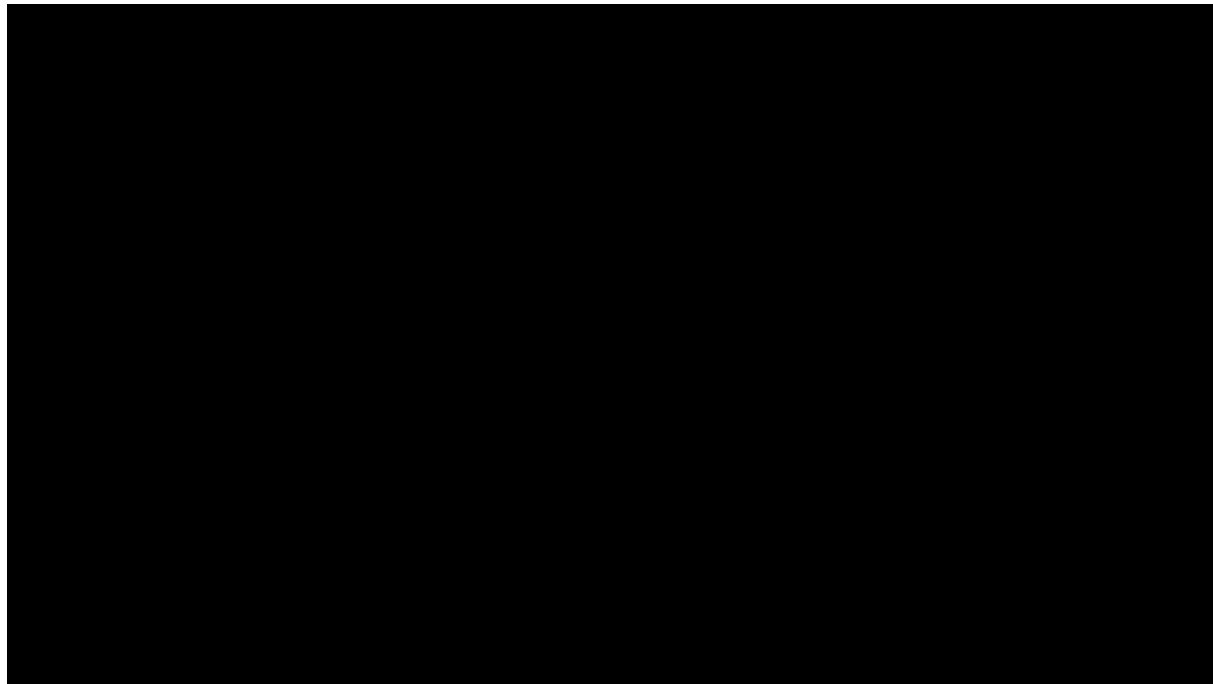
RUHR  
UNIVERSITÄT  
BOCHUM

RUB



**Rome Reborn**

Rekonstruktion und virtuelle Animation des alten Roms

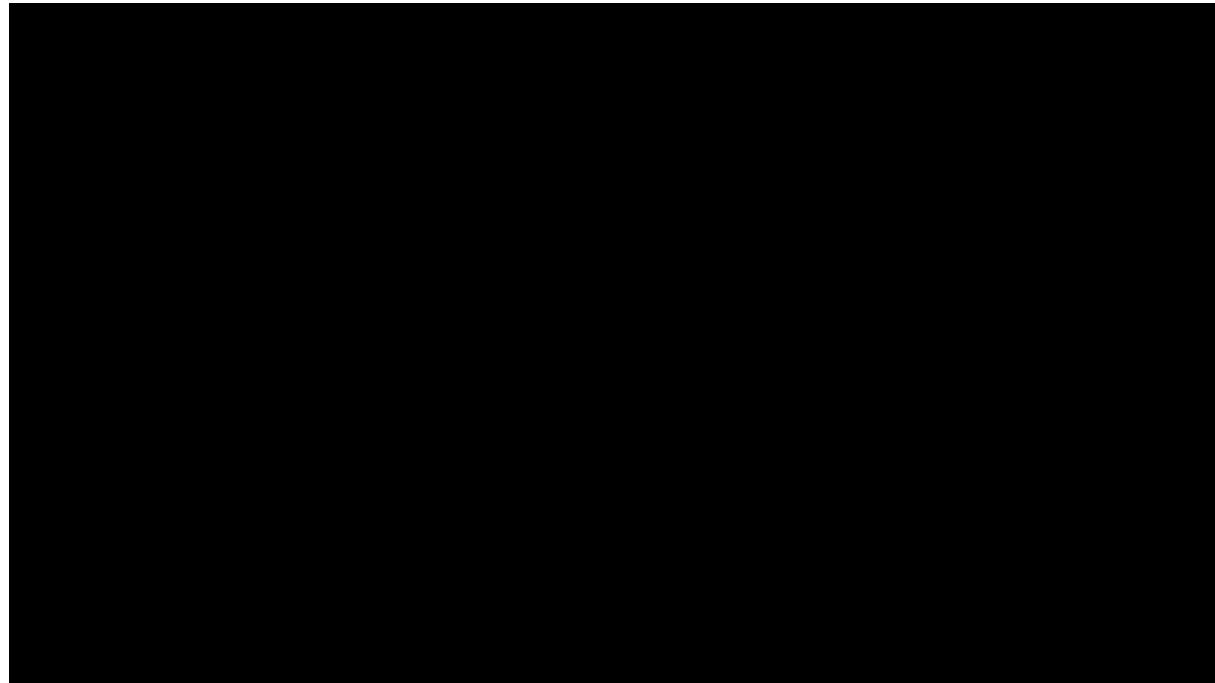


# VISUALISIERUNG GEGENSTAND DER VISUALISIERUNG



## Minecraft Palmyra

Nachbildung der antiken Stadt Palmyra in Minecraft (Universität Leiden)



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

# GEOGRAPHICAL INFORMATION SYSTEMS (GIS)

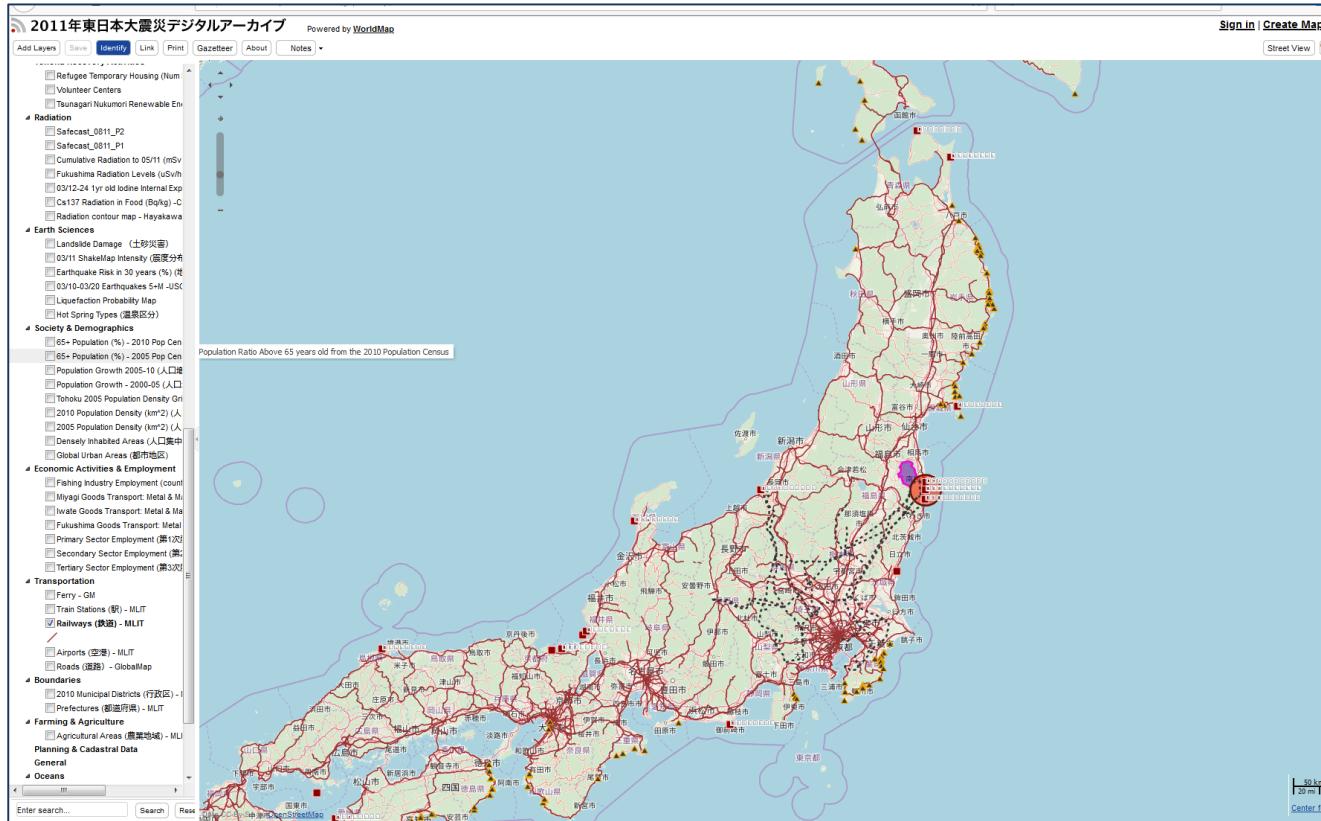
# VISUALISIERUNG

## TYPEN DER VISUALISIERUNG | BEISPIELE



### Harvard World Maps

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



VISUALISIERUNG  
VON DATEN

COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN.

RUHR  
UNIVERSITÄT  
BOCHUM

RUB

# INPUT HARVARD WORLDMAP -> DATA



**HARVARD**  
Dataverse

**G. W. Skinner Data Archive**

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](#)

Search this 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

[Key to Skinner Dataset Names and Abbreviations](#)  
Henderson, Mark; Skinner, G. W., 2018, "Key to Skinner Dataset Names and Abbreviations", <https://doi.org/10.7910/DVN/T7J5DA>, Harvard Dataverse, V1  
Abbreviations for datasets with descriptions of the contents, time periods, and areas covered by the Skinner datasets.

[ChinaT \(1990\) Central Place Analysis](#)  
May 10, 2018 - China - Modern Dataverse

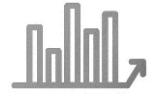
COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN.

RUHR  
UNIVERSITÄT  
BOCHUM

RUB

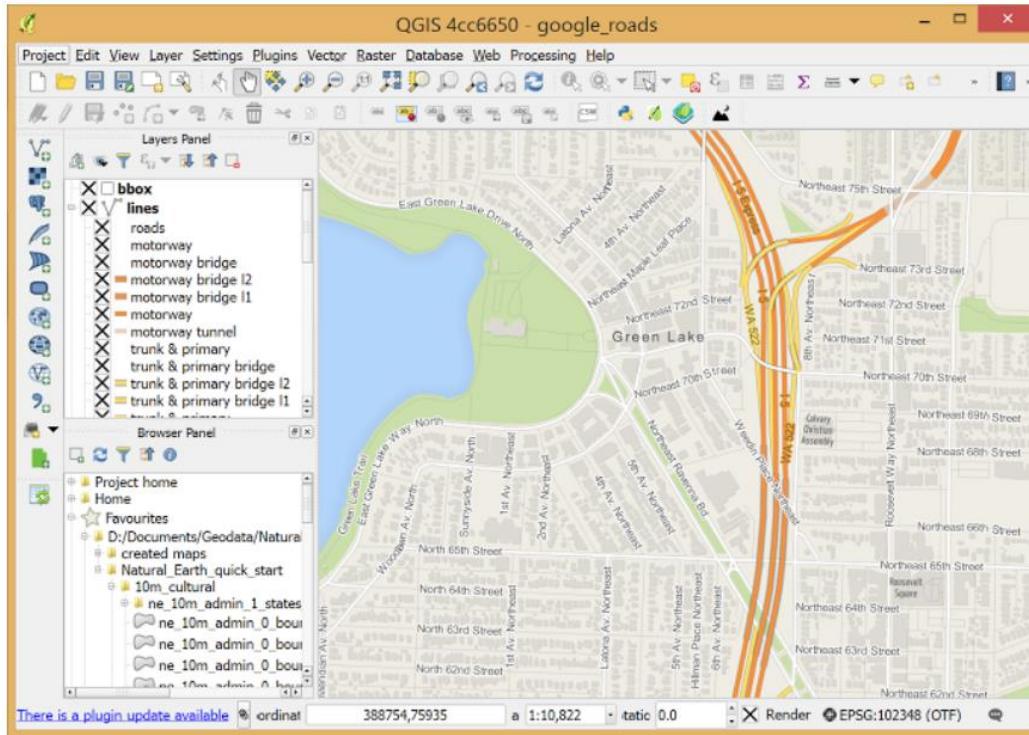
# VISUALISIERUNG

## TYPEN DER VISUALISIERUNG | BEISPIELE



### QGIS

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



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

## VISUALISIERUNG VON DATEN

COMPUTERGESTÜTZTE METHODEN IN DEN OSTASIENWISSENSCHAFTEN.



Markup meets GIS - Visualizing the 'Biographies of Eminent Buddhist Monks'



Figure 6: Buddhist activity in Shandong Province in the 5th century from the *Liang Gaoseng Zhuan*

The aims of this project are:

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\*

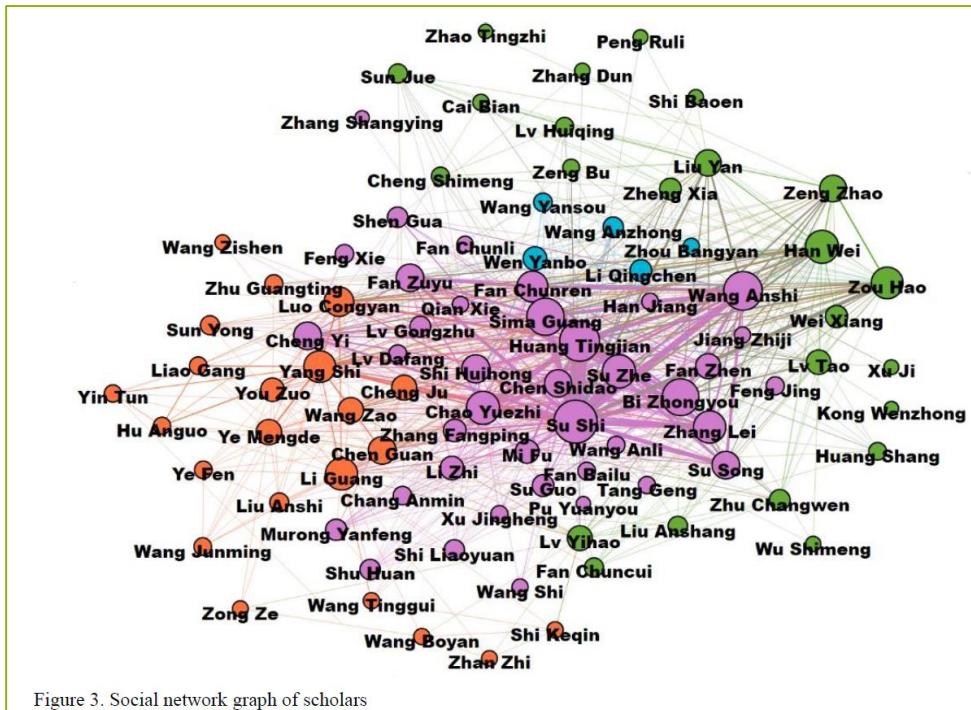


Figure 3. Social network graph of scholars

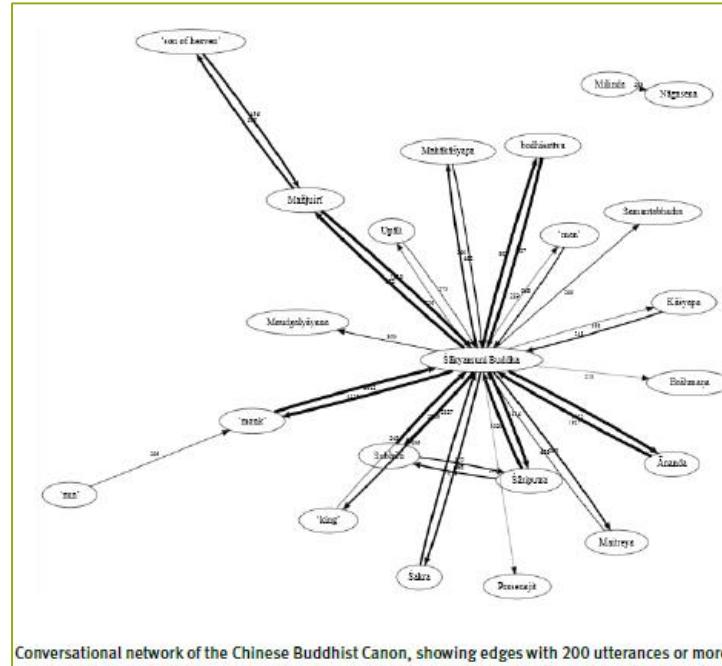


## BEISPIEL: CONVERSATIONAL NETWORK-ANALYSIS

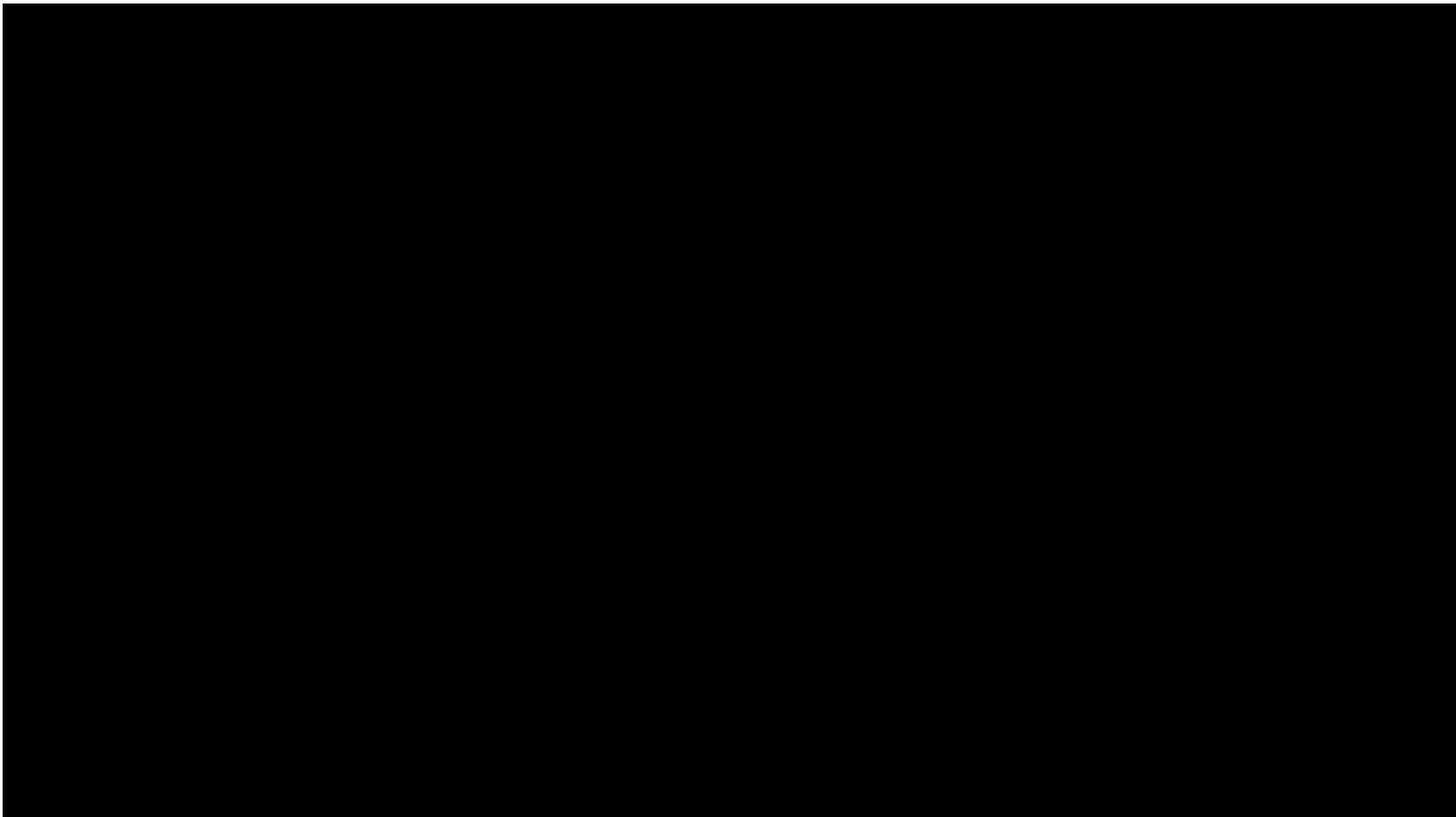
John Lee\*, Tak-sum Wong

### Conversational Network in the Chinese Buddhist Canon

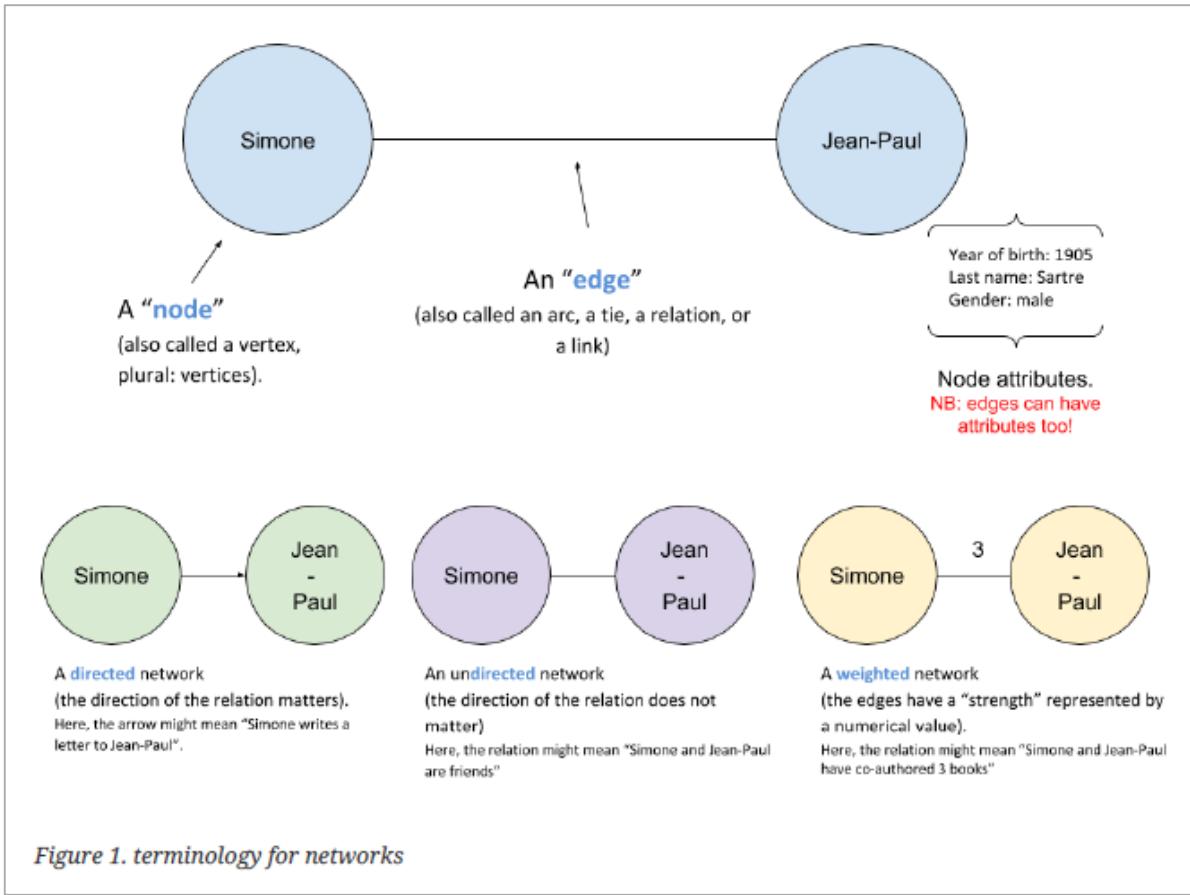
**Abstract:** This article describes a method to analyze characters in a literary text by considering their verbal interactions. This method exploits techniques from computational linguistics to extract all direct speech from a treebank, and to build a conversational network that visualizes the speakers, the listeners and their degree of interaction. We apply this method to create and visualize a conversational network for the Chinese Buddhist Canon. We analyze the protagonists and their interlocutors, and report statistics on their number of utterances and types of listeners, how their speech was reported, and subcommunities in the network.



INPUT  
NETWORK-ANALYSIS | THEORIE



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



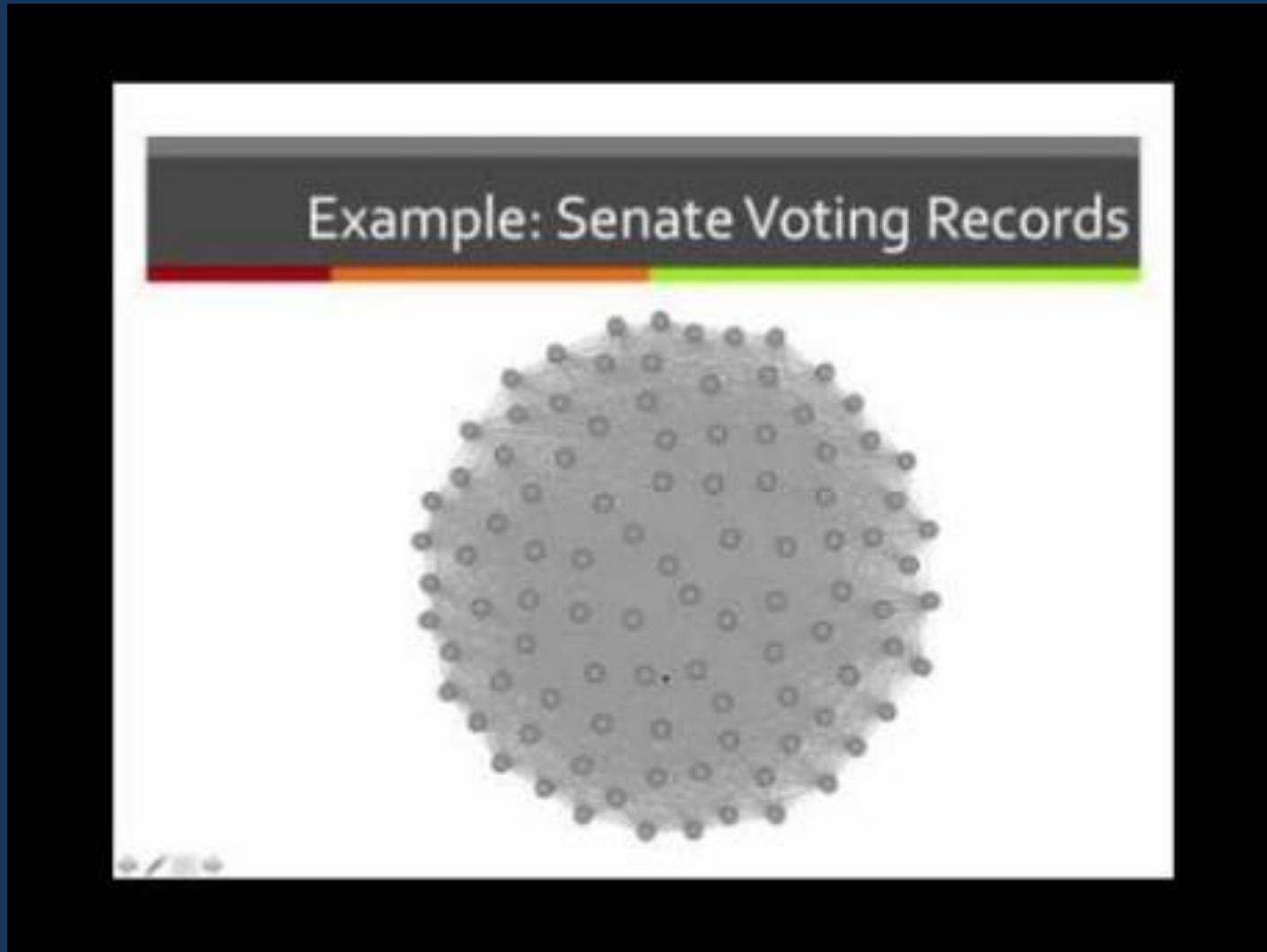
Quelle: „Simple Gephi Project from A to Z“ (2017)

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

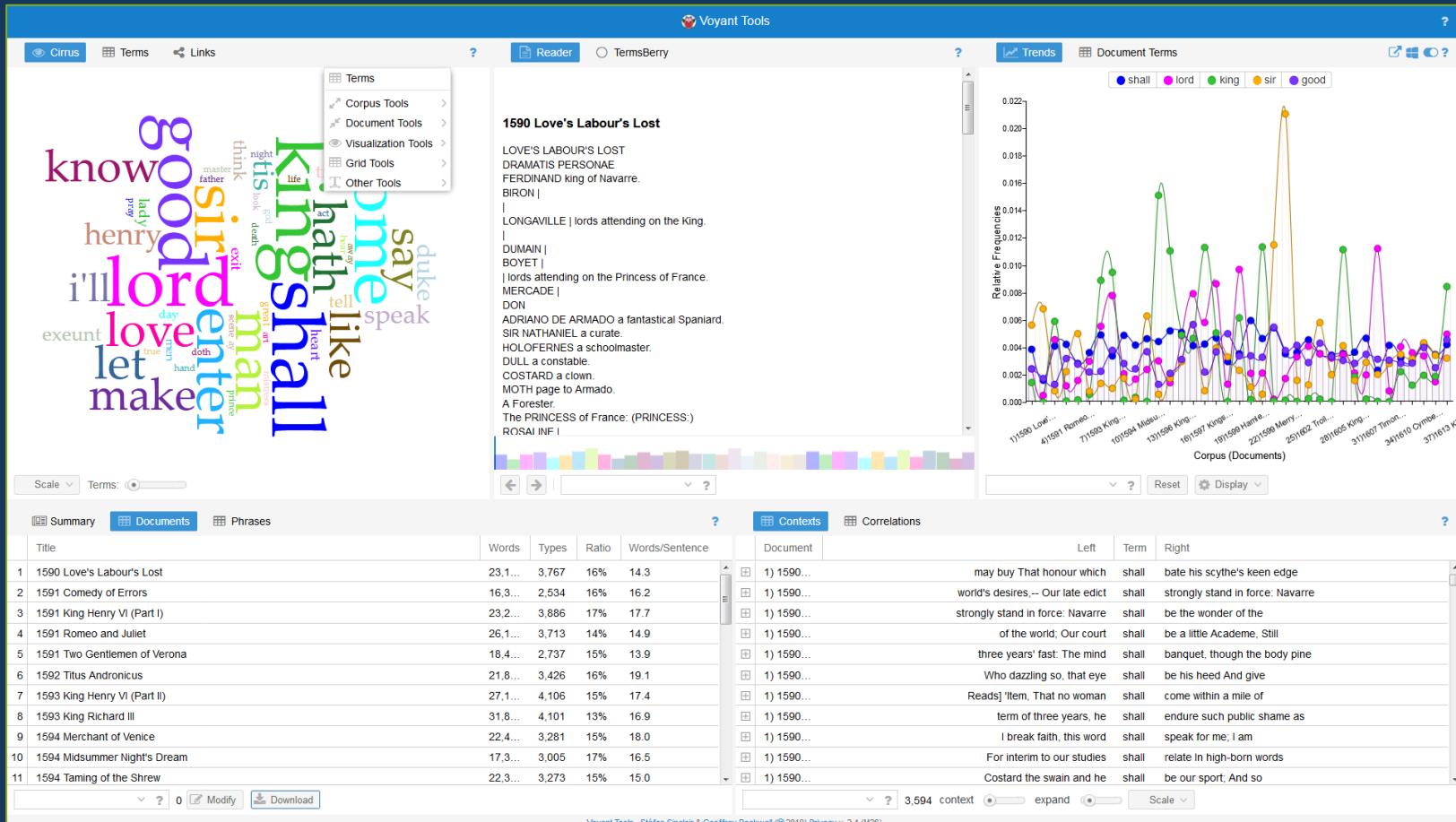
- **Degree**
- **Density**
- **Centrality**
- ✓ **Closeness centrality**
- ✓ **Degree centrality**
- ✓ **Betweenness centrality**



- BEISPIEL 1 | GEPHI UND NETZWERKANALYSE
- BEISPIEL 2 | VOSVIEWER UND BIBLIOMETRISCHE NETZWERKANALYSE
- BEISPIEL 3 | CTEXT-TEXTTOOLS, TEXT-REUSE & NETZWERKANALYSE



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





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



Voyant Tools interface showing various tools and a detailed analysis of "1590 Love's Labour's Lost".

**Tools Panel:** Cirrus, Terms, Links, Reader, TermsBerry, Trends, Document Terms.

**Visualizations:** A word cloud on the left and a line graph on the right showing relative frequencies of words like "shall", "lord", "king", "sir", and "good" across different documents.

**Text Analysis:** The main panel displays the text of "1590 Love's Labour's Lost" with annotations. Below it is a bar chart of word counts per document.

**Tables:** Summary, Documents, Phrases, Contexts, Correlations.

**Download:** A button at the bottom left labeled "Download" is circled in red.



TermsBerry

Terms

	Term	Count ↓	Relative	P...	Trend
1	small	3594	4,012	0.7	
2	lord	3346	3,735	1.4	
3	king	3309	3,694	0.1	
4	sir	3031	3,384	13.6	
5	good	2883	3,219	0.5	
6	come	2569	2,868	0.4	
7	enter	2403	2,683	0.4	
8	let	2108	2,353	-0.2	
9	love	1956	2,184	4.7	
10	hath	1913	2,136	-0.1	
24,807					



- [Reader](#)
- [Summary](#)
- [Cirrus](#)
- [Terms](#)
- [Trends](#)
- [Bubblelines](#)
- [Document Terms](#)
- [Contexts](#)
- [Collocates](#)
- [Correlations \(?\)](#)
- [Phrases](#)



## CORPUS TOOLS → CIRRUS

Voyant Tools

Cirrus

Document Terms

Scale Terms: 0

Scale Terms: 0

Scale Terms: 0

**Summary** **Documents** **Phrases**

Distinctive words (compared to the rest of the corpus):

- 1590 Love's Labour's Lost: **biron** (194), **adriano** (112), **boyet** (102), **armado** (121), **ferdinand** (127).
- 1591 Comedy of Errors: **syracuse** (232), **dromio** (221), **antipholus** (219), **ephesus** (169), **adriana** (90).
- 1591 King Henry VI (Part...): **pucelle** (79), **somerset** (57), **reignier** (47), **suffolk** (58), **orleans** (48).
- 1591 Romeo and Juliet: **capulet** (143), **mercutio** (84), **tybalt** (69), **laurence** (73), **frial** (99).
- 1591 Two Gentlemen of...: **proteus** (220), **julia** (144), **silvia** (125), **valentine** (216), **launce** (82).
- 1592 Titus Andronicus: **andronicus** (249), **titus** (200), **aaron** (97), **tamora** (91), **lavinia** (76).
- 1593 King Henry VI (Part...): **cade** (98), **suffolk** (139), **york** (135), **margaret** (94), **warwick** (78).
- 1593 King Richard III: **buckingham** (152), **elizabeth** (114), **catesby** (75), **richard** (230), **hastings** (96).

Contexts Correlations

Document	Left	Term	Right
1) 1590 ...	may buy That honour which	shall	bate his scythe's keen edge
1) 1590 ...	world's desires,-- Our late edict	shall	strongly stand in force: Navarre
1) 1590 ...	strongly stand in force: Navarre	shall	be the wonder of the
1) 1590 ...	of the world; Our court	shall	be a little Academe, Still
1) 1590 ...	three years' fast: The mind	shall	banquet, though the body pine

3,594 contexts Scale

Voyant Tools , Stéfan Sinclair & Geoffrey Rockwell (© 2019) Privacy v. 2.4 (M26)



## CORPUS TOOLS → TERMS &amp; TRENDS

Voyant Tools

Cirrus Terms Links

TermsBerry Terms Trends Document Terms

Relative Frequencies

Relative Frequencies

Summary Documents Phrases Contexts Correlations

This corpus has 37 documents with 895,737 total words and 25,114 unique word forms. Created about a month ago.

Document Length:

- Longest: 1599 Hamlet (32396); 1593 King Richard III (31846); 1608 Coriolanus (29485); 1610 Cymbeline (29204); 1597 King Henry IV (Part... (28225)
- Shortest: 1591 Comedy of Errors (16339); 1594 Midsummer Night's... (17332); 1611 Tempest (17582); 1605 Macbeth (18414); 1591 Two Gentlemen of... (18443)

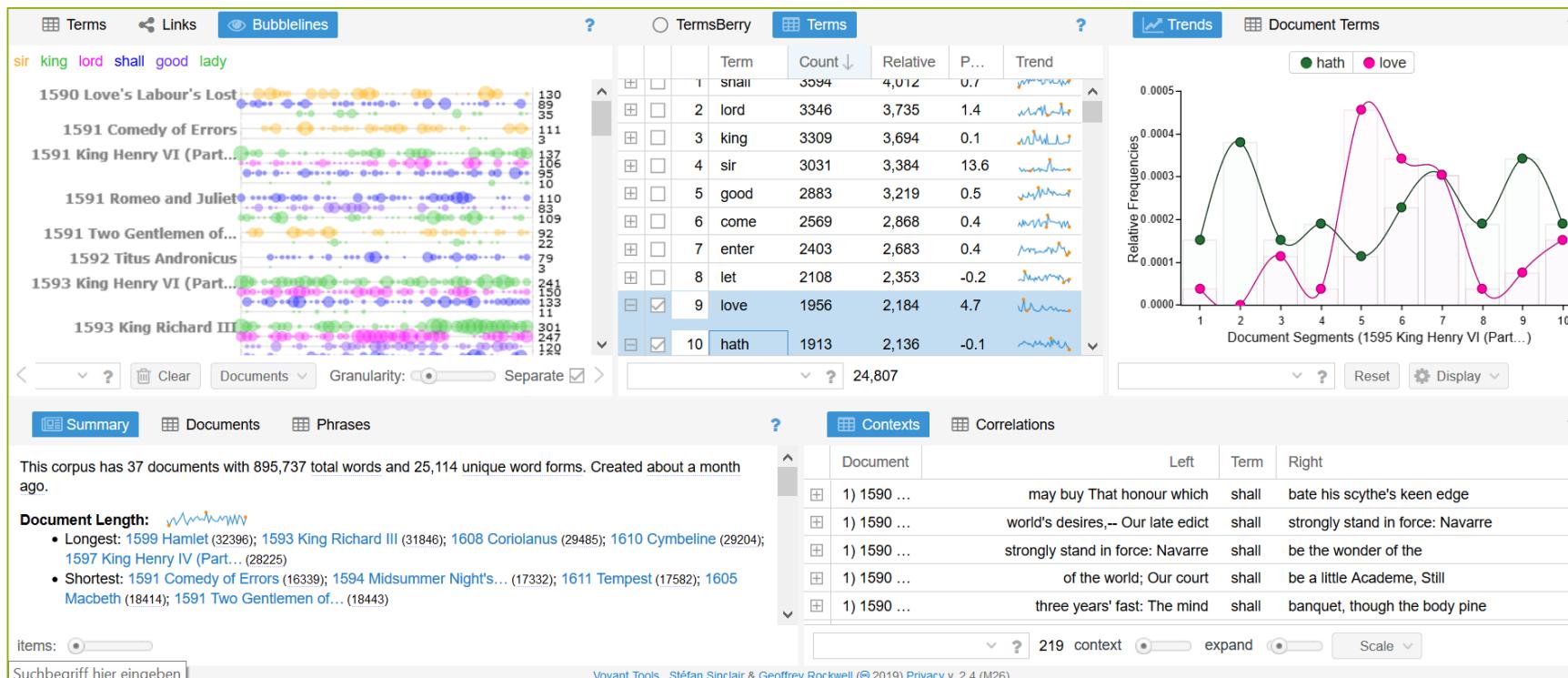
Items: Scale

Voyant Tools - Stéfan Sinclair & Geoffrey Rockwell (© 2019) Privacy v. 2.4 (M26)

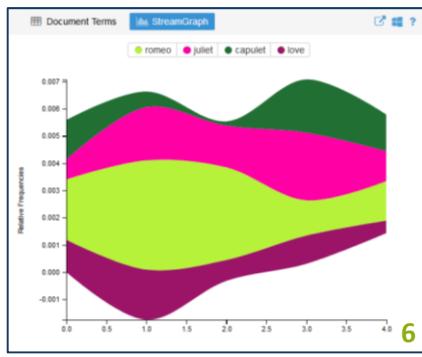
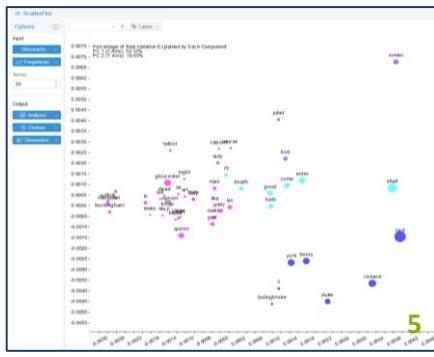
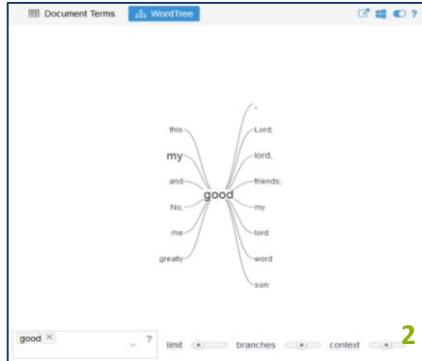
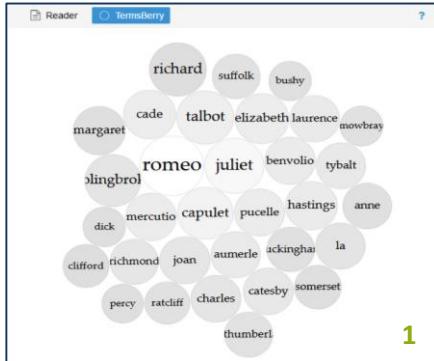
Quelle: <https://voyant-tools.org/?panels=cirrus%2Ccorpusterms%2Ctrends%2Csummary%2Ccontexts&corpus=shakespeare>



## CORPUS TOOLS → BUBBLE LINES



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



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 stopword lists in various languages or create your own. For tools that support stopword 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](https://qgis.org/de/site/Open%20Source)
- **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](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](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](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](http://ctext.org)
- **Visual Linguistics Project:** [https://www.bubenhofer.com/graphics/bubenhofer\\_visual\\_linguistik.mp4](https://www.bubenhofer.com/graphics/bubenhofer_visual_linguistik.mp4)
- **GeoCollocations:** <http://www.bubenhofer.com/geocollocations/>
- **RUB Zentrum für Geoinformation:** <http://www.izg.rub.de/>