

# TextVIS for DH

Florian Windhager, Asil Çetin-Aufricht

# Outline

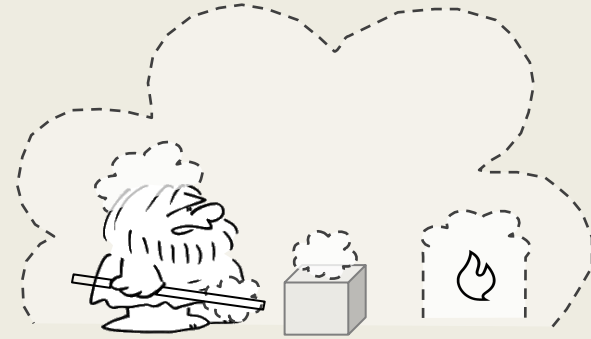
1. Introduction
2. Surveys
3. Tools & Examples
4. Live Exercise with Voyant
5. Discussion & Challenges

# Humanities

*Those branches of knowledge that concern themselves with human beings and their culture or with analytic and critical methods of inquiry derived from an appreciation of human values and of the unique ability of the human spirit to express itself.*

<https://www.britannica.com/topic/humanities>

## Concepts of “Culture”



a) artful / aesthetic / unique / sacred things

b) useful / practical / daily / mundane  
things

+ c) awful / unintended / discomforting things



Maps: <https://www.google.de/maps/>



*„The most important issues today  
are not photographable“*

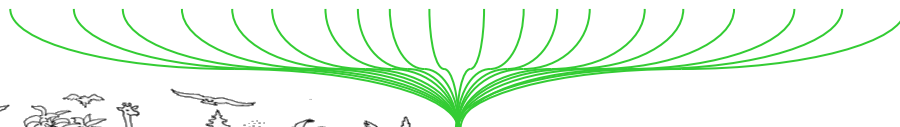
Moritz Stefaner

Texts

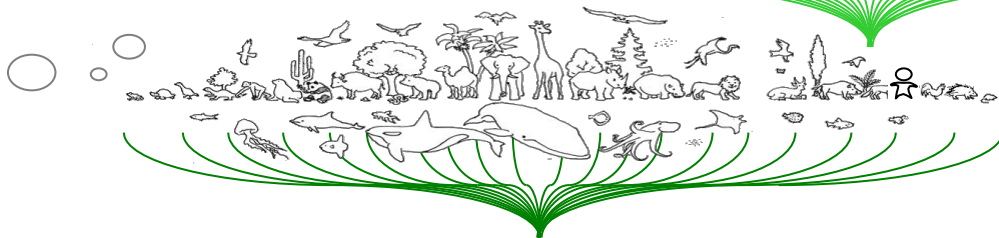
Cultural objects & artifacts



“Cultural Materials”



“Culture”

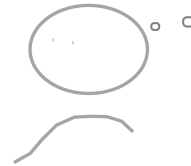
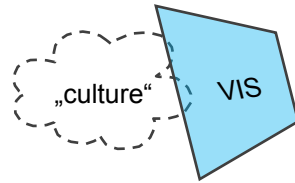


“Nature”

„World“

- **different types & aggregations of „cultural data“**

- texts & intertexts
- images
- objects
- music
- movies
- performances
- customs & practices
- etc.

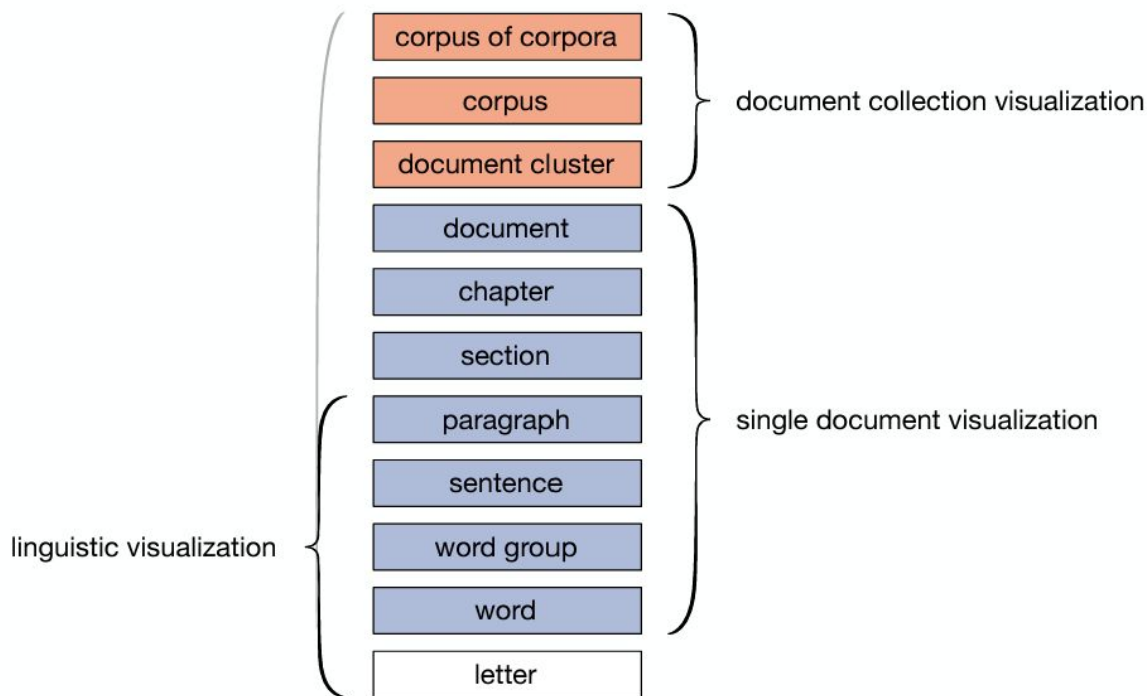


- object **collections**, corpora, editions & aggregated collections

- a) **expert users** (e.g. arts & humanities scholars, educators, cultural journalists, tourist guides, etc.)
- b) **casual users** (everyday visitors, non-experts and non-professionals with leisure time access)

- to see / observe
- to remember
- to represent
- to mediate
- to *analyze*
- to enjoy / to savour
  
- to interpret
- to evaluate
- to contextualize
- *to understand*

# Text Units Hierarchy



## *How do we study texts?*

### **„Close Reading“**

*'Essentially, close reading means reading to uncover layers of meaning that lead to deep comprehension'.  
(Boyles, 2013)*

*In other words, close reading is the thorough interpretation of a text passage by the determination of central themes and the analysis of their development.*

*Moreover, close reading includes the analysis of (1) individuals, events and ideas, their development and interaction; (2) used words and phrases; (3) text structure and style; and (4) argument patterns  
(Jasinski, 2001).*

*(Jänicke et al. 2016)*

## 3 Phases of Close Reading

### Read for 3 purposes

#### Phase 1

WHAT the text says



##### Read for:

- who, what, when, where
- main idea

#### Phase 2

HOW the text says it



##### Zoom in for:

- new words
- vocabulary
- point of view
- author's purpose

#### Phase 3

What the text MEANS



##### Make connections for:

- text-to-text ideas
- evidence
- showing your thinking

# CLOSE READING STEPS



#### FIRST READING

- Read Silently
- Think
- Question
- Annotate
- Retell
- Summarize
- Discuss with partner



#### SECOND READING

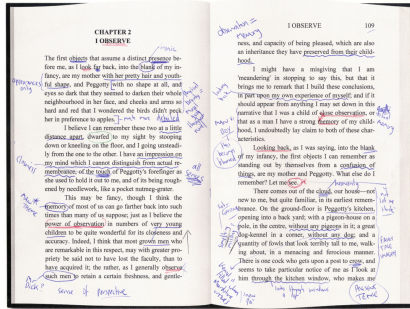
- Teacher reads aloud
- Student listens, thinks, writes
  - structure
  - text features
  - context clues
  - author's purpose
  - point of view
  - word meaning

#### THIRD READING

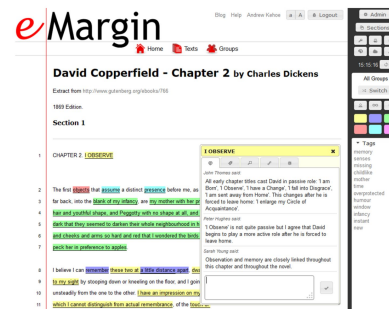
- Re-read
- Make connections
- Compare/contrast
- Analyze
- Evaluate
- Find evidence
- Write/cite evidence

Aligns with Common Core State Standards for Reading 2-8

Copyright © 2014 abcteach.com

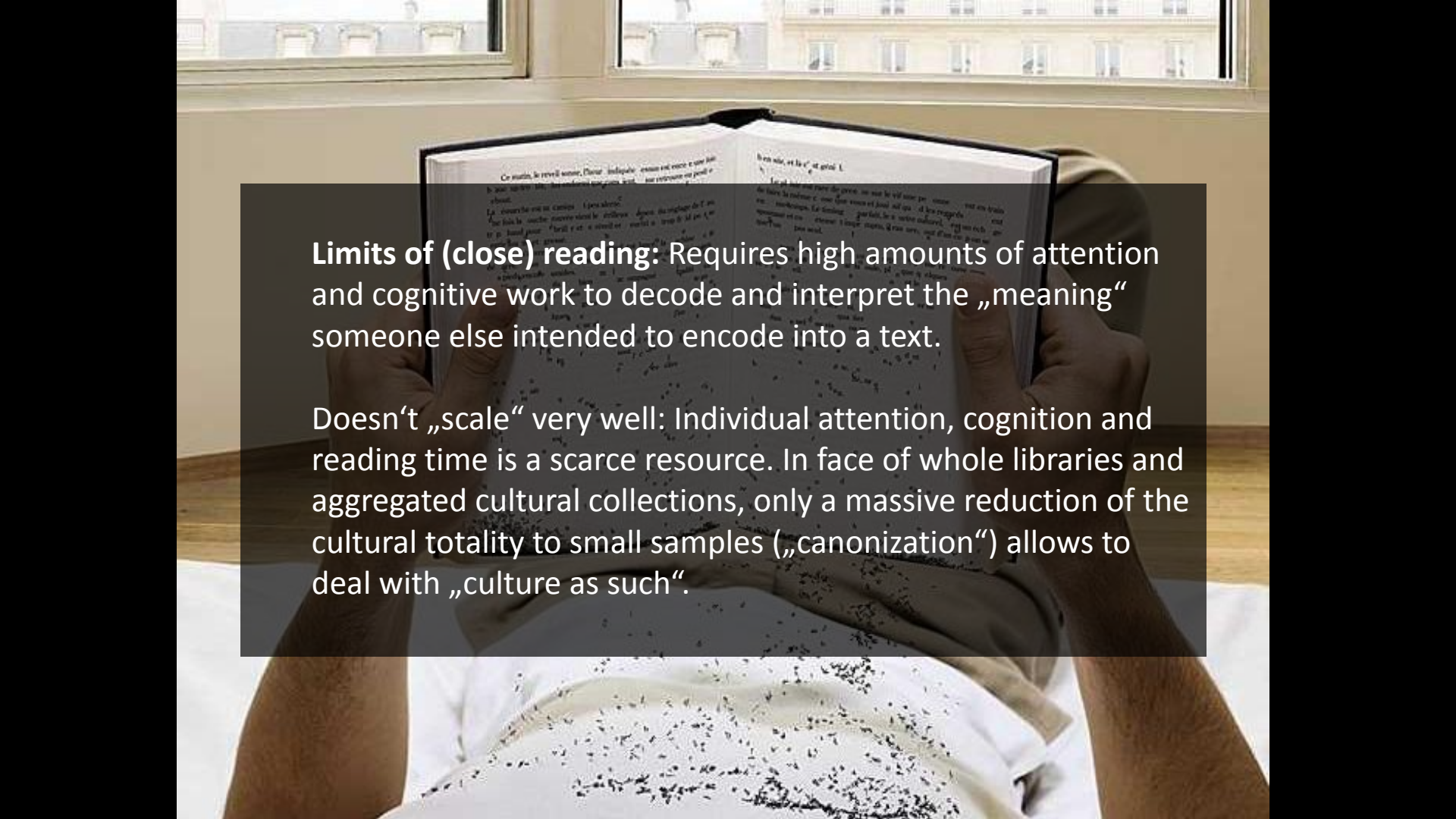


Traditional practice of  
annotation + visualization



Digital practice of  
annotation / NLP + visualization

(Jänicke et al., 2017, [link](#))

A person is sitting in a room, reading an open book. The book is held open with both hands. The pages contain text in French. The person is wearing a light-colored long-sleeved shirt. In the background, there is a window with a view of a building with many windows. The text is overlaid on a semi-transparent dark grey box.

**Limits of (close) reading:** Requires high amounts of attention and cognitive work to decode and interpret the „meaning“ someone else intended to encode into a text.

Doesn't „scale“ very well: Individual attention, cognition and reading time is a scarce resource. In face of whole libraries and aggregated cultural collections, only a massive reduction of the cultural totality to small samples („canonization“) allows to deal with „culture as such“.

# “Distant reading” (Moretti, 2000)

“Ars longa,  
vita brevis.”



1) Art is long .....

2) life is short ....

3) comp. is fast ...

4) visualizations make big data structures & dynamics visible .....

## How do we study texts?

### „Close Reading“

*'Essentially, close reading means reading to uncover layers of meaning that lead to deep comprehension'. (Boyles, 2013)*

*In other words, close reading is the thorough interpretation of a text passage by the determination of central themes and the analysis of their development.*

*Moreover, close reading includes the analysis of (1) individuals, events and ideas, their development and interaction; (2) used words and phrases; (3) text structure and style; and (4) argument patterns (Jasinski, 2001).*

### „Distant Reading“

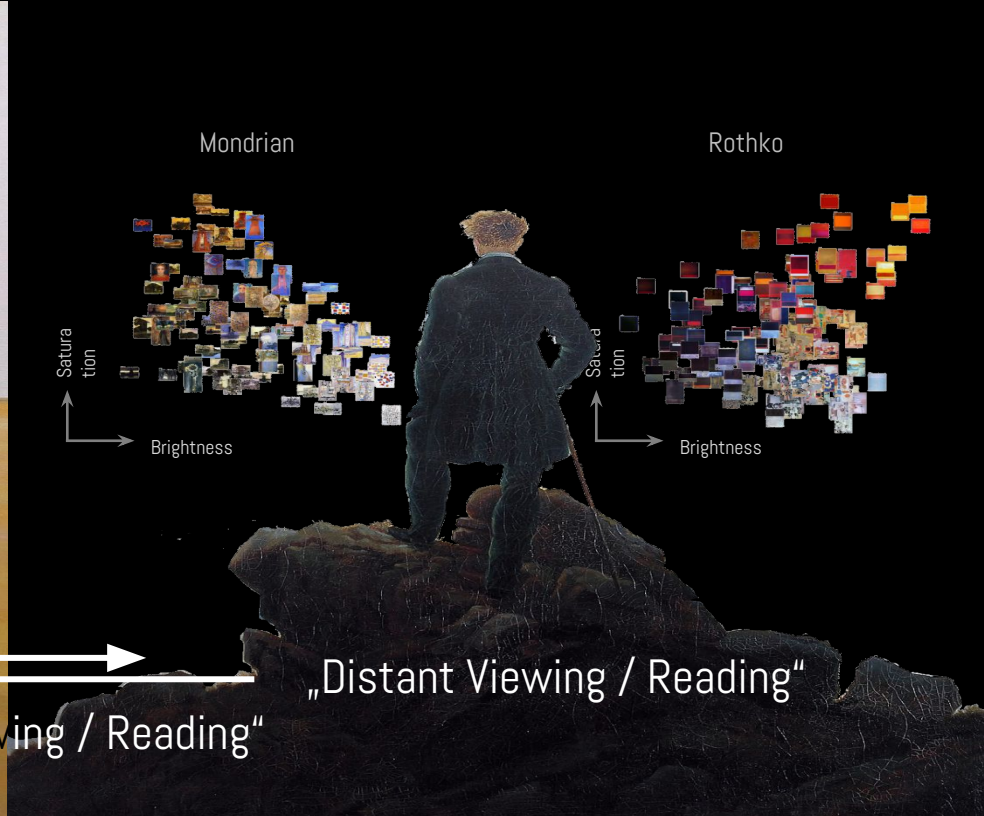
*'While close reading retains the ability to read the source text without dissolving its structure, distant reading does the exact opposite. It aims to generate an abstract view by shifting from observing textual content to visualizing global features of a single or of multiple text(s).*

*Moretti [Moretti, 2013] describes distant reading as 'a little pact with the devil: we know how to read texts, now let's learn how not to read them'. In 2005, he introduces his idea of distant reading (Moretti, 2005) with three examples using:*

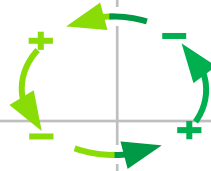
- *graphs to analyse genre change of historical novels,*
- *maps to illustrate geographical aspects of novels and*
- *trees to classify different types of detective stories.*

*(Jänicke et al. 2016)*



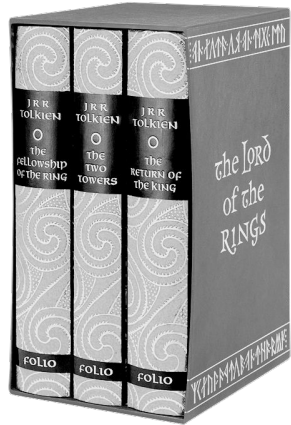


| Distant Reading & Viewing   | Close Reading & Viewing   |
|---|---|
| <p><b>Stärken:</b> Erfassung und Analyse von kulturellen Objekten im großen Maßstab ("distant reading &amp; viewing")</p> | <p><b>Limitationen:</b> Fokus on relativ wenige Objekte, aufgrund beschränkter Zeit, Aufmerksamkeit, und Verfügbarkeit</p>                          |
| <p><b>Limitationen:</b> Analyse möglich nur für relative einfache Merkmale und Eigenschaften von kulturellen Objekten</p> | <p><b>Stärken:</b> Historisch-kritische Analyse, Interpretation und Kontextualisierung von kulturellen Objekten ("close reading &amp; viewing")</p> |

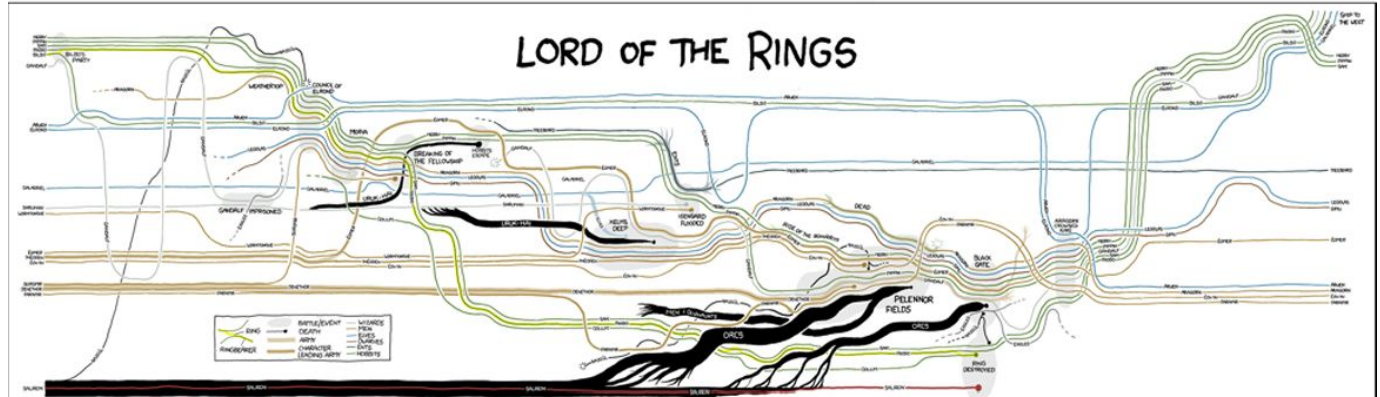


A few examples...

A fantasy trilogy...



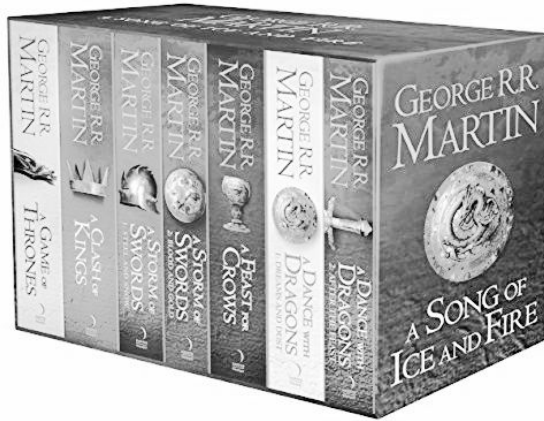
THESE CHARTS SHOW MOVIE CHARACTER INTERACTIONS.  
THE HORIZONTAL AXIS IS TIME. THE VERTICAL GROUPING OF THE  
LINES INDICATES WHICH CHARACTERS ARE TOGETHER AT A GIVEN TIME.



<https://xkcd.com/657/>

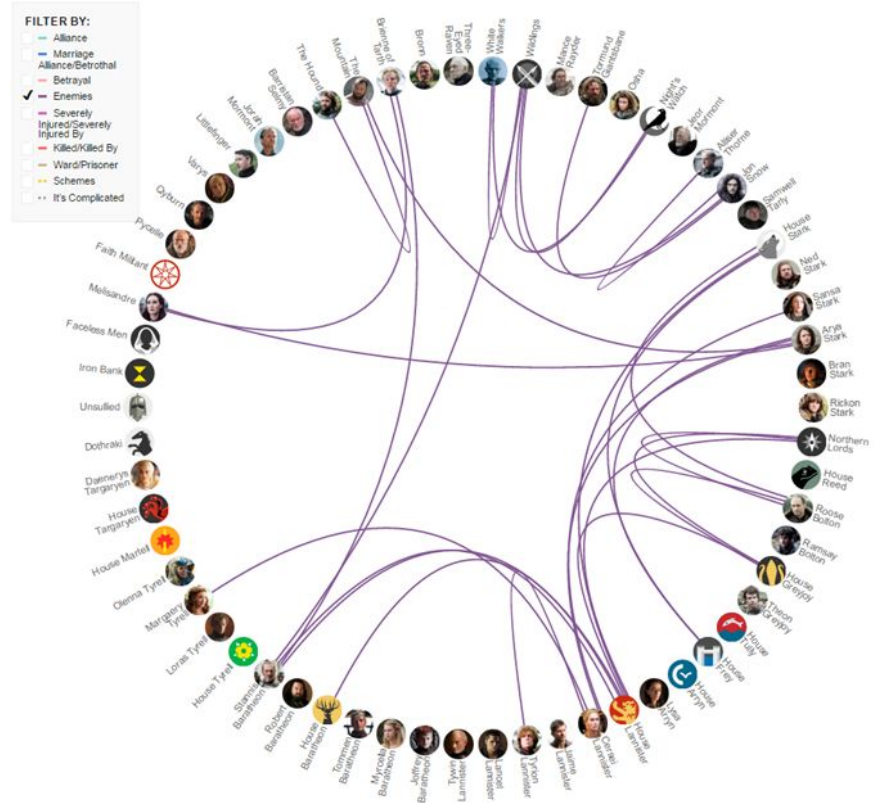
<http://www.ycwu.org/projects/infovis13.html>

# A fantasy heptalogy... (?)



Season 1, Beginning    Season 2, Beginning    Season 3, Beginning    Season 4, Beginning    **Season 5, Beginning**    Mid-Season 5

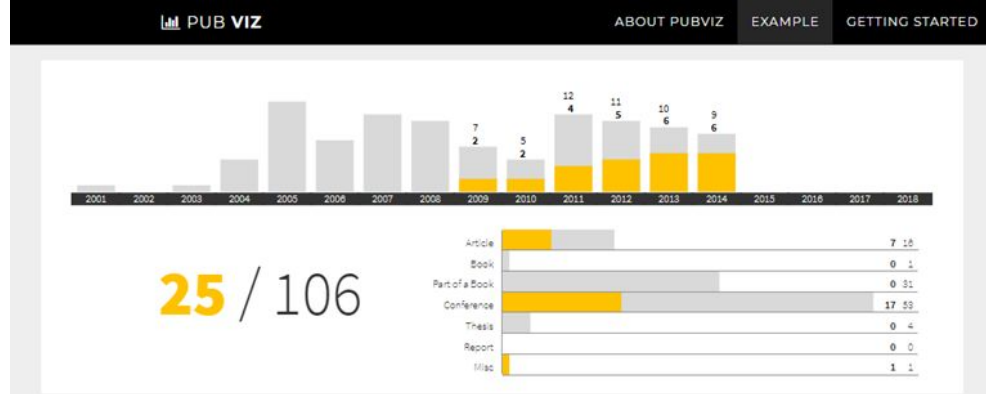
Hover over a character to see their connections. Click on a character for more detail.



# The output of a research group

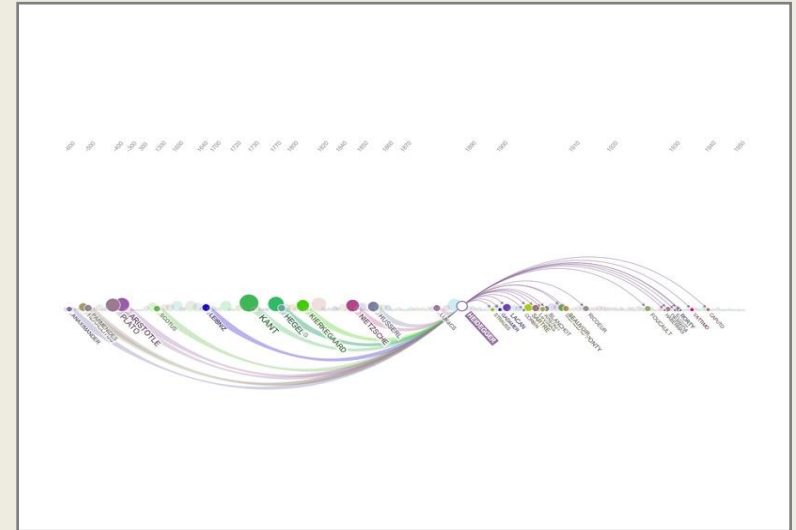
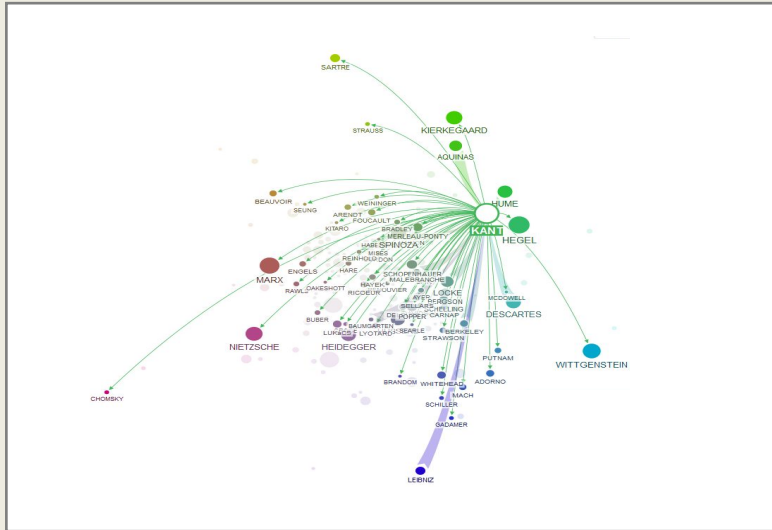


<http://pubviz.fhstp.ac.at/>



| Year | Type       | Publication Title  | Year |
|------|------------|--|------|
| 2014 | Article    | Mind the Time: Unleashing Temporal Aspects in Pattern Discovery  | 2014 |
| 2014 | Conference | Problem Characterization and Abstraction for Visual Analytics in Behavior-Based Malware Pattern Analysis | 2014 |
| 2014 | Conference | Qualizon Graphs: Space-Efficient Time-Series Visualization with Qualitative Abstractions                 | 2014 |
| 2014 | Conference | Showing Important Facts to a Critical Audience by Means Beyond Desktop Computing                         | 2014 |
| 2014 | Conference | User Tasks for Evaluation: Unangling the Terminology Throughout Visualization Design and Development     | 2014 |

# The Network of Philosophy



## 2. Surveys

# Text Visualization Techniques: Taxonomy, Visual Survey, and Community Insights

Kostiantyn Kucher\*

Andreas Kerren†

ISOVIS Group, Department of Computer Science, Linnaeus University, Växjö, Sweden



Figure 1: The web-based user interface of our visual survey called *Text Visualization Browser*. By using the interaction panel on the left hand side, researchers can look for specific visualization techniques and filter out entries with respect to a set of categories (cf. the taxonomy given in Sect. 3). Details for a selected entry are shown by clicking on a thumbnail image in the main view. The survey contains 141 categorized visualization techniques by January 19, 2015.

## ABSTRACT

Text visualization has become a growing and increasingly important subfield of information visualization. Thus, it is getting harder for researchers to look for related work with specific tasks or visual metaphors in mind. In this paper, we present an interactive visual survey of text visualization techniques that can be used for the purposes of search for related work, introduction to the subfield and gaining insight into research trends. We describe the taxonomy used for categorization of text visualization techniques and compare it to approaches employed in several other surveys. Finally, we present results of analyses performed on the entries data.

**Keywords:** Visualization, text visualization, survey, interaction, web-based systems, taxonomy, community analysis

in other cases on results of text mining algorithms. In the same way, they can be rather general or very specialized and dedicated to specific analytic tasks or application domains. This is the reason why we have decided to construct a taxonomy with numerous categories and subcategories that is exploited by the survey browser in order to facilitate the interactive exploration of the survey entries. Our visual survey has been implemented as a web page and includes 141 techniques at present. We present peer-reviewed work in InfoVis, Visual Analytics and related research fields. After a short discussion on relevant surveys in the following, we highlight the taxonomy used by our survey browser, some implementation details, and the results of analyses conducted on the collected entries data. The present paper is based on our previous poster abstract [20].

(Kucher & Kerren, 2015)

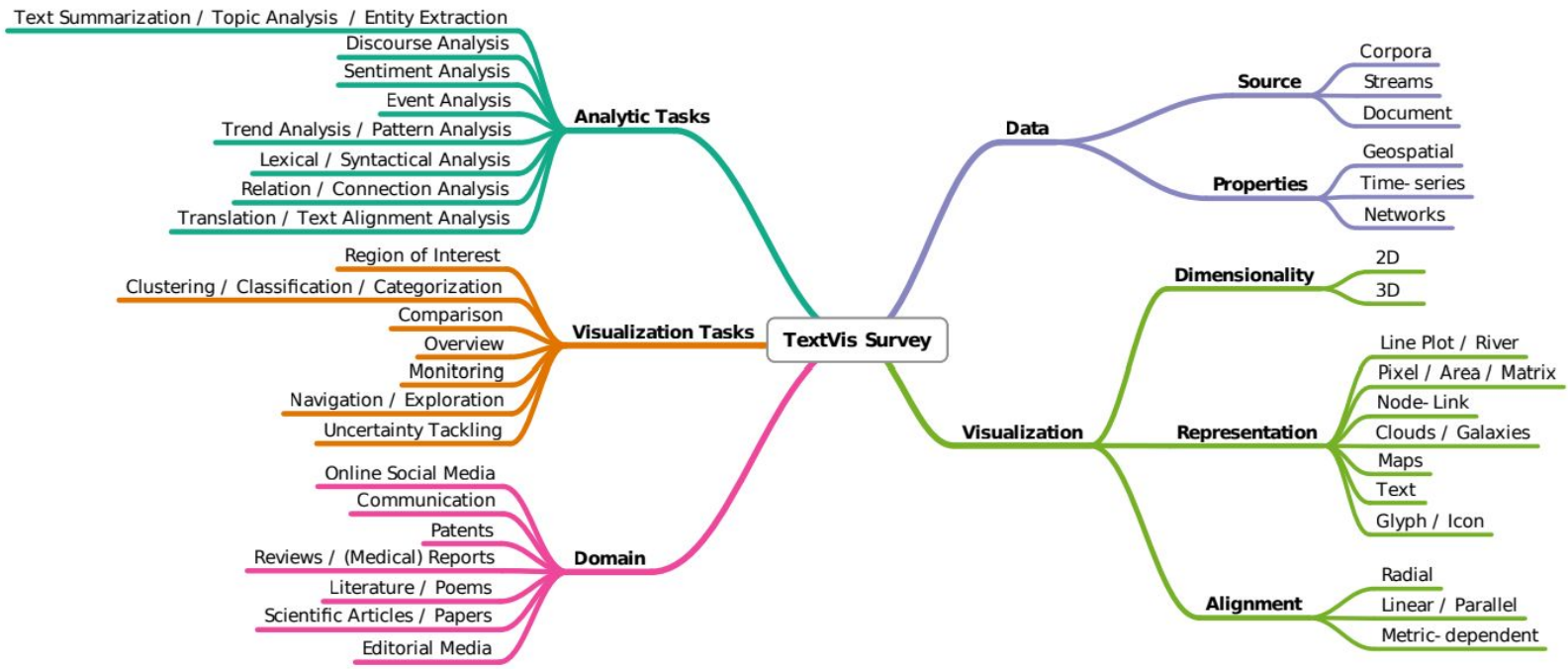


Figure 2: The taxonomy of text visualization techniques used in our visual survey. We focus on the description on the left hand side of the figure in this paper and only briefly summarize the right side in Subsect. 3.4 which should be self-evident for the visualization community.

# Text Visualization Browser

A Visual Survey of Text Visualization Techniques (IEEE PacificVis 2015 short paper)

Provided by ISOVIS group

[About](#) [Summary](#) [Add entry](#) [Other surveys](#) ▾

Techniques displayed:

440

Search:

Time filter:

1976

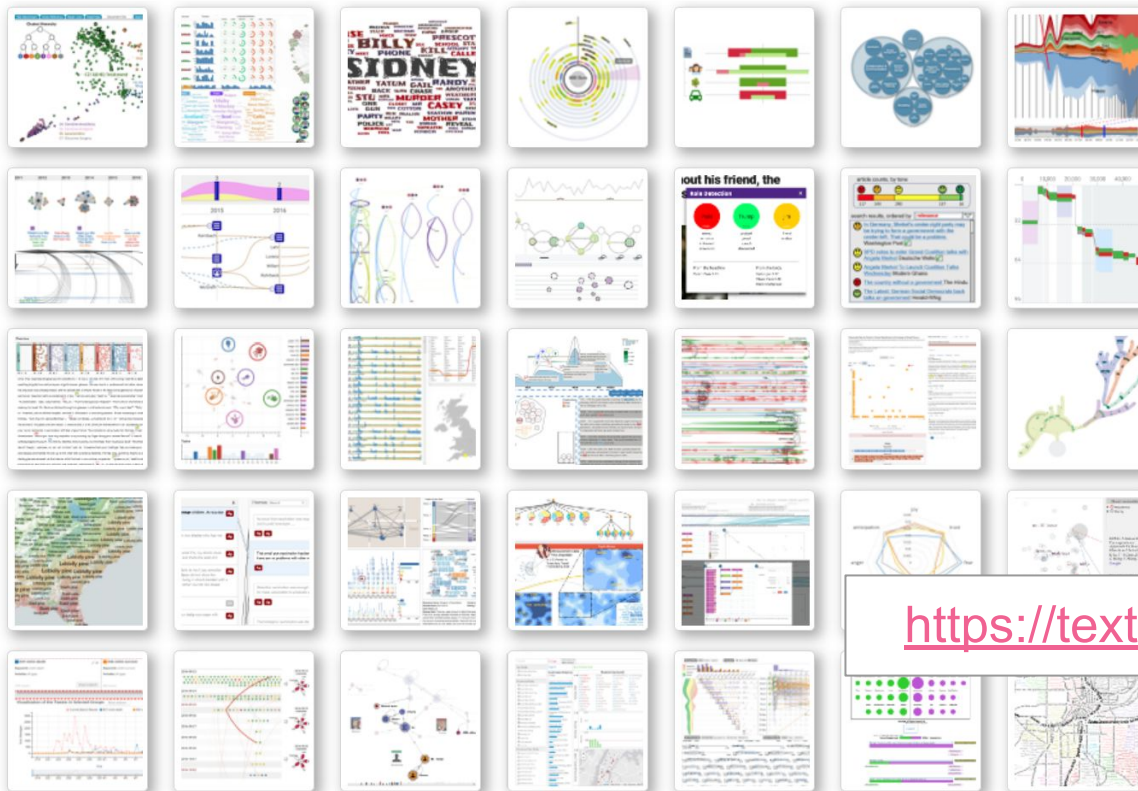
2019



Analytic Tasks



Visualization Tasks



<https://textvis.lnu.se/>



## Visual Text Analysis in Digital Humanities

S. Jänicke<sup>1</sup>, G. Franzini<sup>2</sup>, M. F. Cheema<sup>1</sup> and G. Scheuermann<sup>1</sup>

<sup>1</sup>Image and Signal Processing Group, Department of Computer Science, Leipzig University, Germany  
{stjaenicke, faisal, scheuermann}@informatik.uni-leipzig.de

<sup>2</sup>Göttingen Centre for Digital Humanities, University of Göttingen, Germany  
gfranzini@gwdg.de

---

### Abstract

*In 2005, Franco Moretti introduced Distant Reading to analyse entire literary text collections. This was a rather revolutionary idea compared to the traditional Close Reading, which focuses on the thorough interpretation of an individual work. Both reading techniques are the prior means of Visual Text Analysis. We present an overview of the research conducted since 2005 on supporting text analysis tasks with close and distant reading visualizations in the digital humanities. Therefore, we classify the observed papers according to a taxonomy of text analysis tasks, categorize applied close and distant reading techniques to support the investigation of these tasks and illustrate approaches that combine both reading techniques in order to provide a multi-faceted view of the textual data. In addition, we take a look at the used text sources and at the typical data transformation steps required for the proposed visualizations. Finally, we summarize collaboration experiences when developing visualizations for close and distant reading, and we give an outlook on future challenges in that research area.*

**Keywords:** digital humanities, survey, visual text analysis, close reading, distant reading

**ACM CCS:** H.5.2 [Information Interfaces and Presentation]: User Interfaces–Evaluation/methodology

---

### 1. Introduction

Traditionally, humanities scholars carrying out research on a specific or on multiple literary work(s) are interested in the analysis of related texts or text passages. But the digital age has opened possibilities for scholars to enhance their traditional workflows. Enabled by digitization projects, humanities scholars can nowadays reach a large number of digitized texts through web portals such as Google Books [Goo15] and Internet Archive [IA15]. Digital editions exist

Developed in the late 1980s [Hoc04], the digital humanities primarily focused on designing standards to represent cultural heritage data such as the Text Encoding Initiative (TEI) for texts [TEI15], and to aggregate, digitize and deliver data. In the last years, visualization techniques have gained more and more importance when it comes to analysing data. For example, Saito [SO10] introduced her 2010 digital humanities conference paper with ‘In recent years, people have tended to be overwhelmed by a vast amount of information in

(Jänicke et al., 2017)

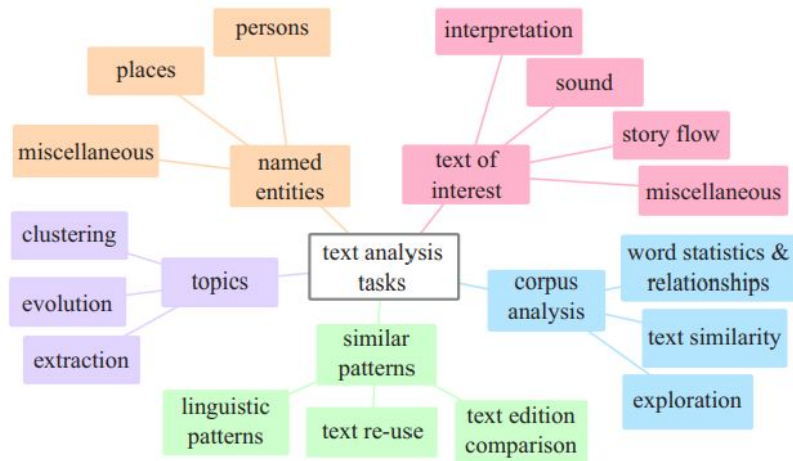


Figure 5: Taxonomy of text analysis tasks.

|                                 | Close Reading |       |           |        |             | Distant Reading |            |      |           |        | Combining     |           |                      |          |
|---------------------------------|---------------|-------|-----------|--------|-------------|-----------------|------------|------|-----------|--------|---------------|-----------|----------------------|----------|
|                                 | Plain         | Color | Font size | Glyphs | Connections | Heat maps       | Tag clouds | Maps | Timelines | Graphs | Miscellaneous | Bottom-up | Top-down & Bottom-up | Top-down |
| <i>named entities</i>           |               |       |           |        |             |                 |            |      |           |        |               |           |                      |          |
| places                          | 1             | 2     |           |        |             | 1               | 12         | 5    | 1         |        |               | 1         |                      | 2        |
| persons                         | 2             |       |           |        |             | 1               | 2          | 1    | 15        | 1      |               | 1         | 1                    |          |
| miscellaneous                   |               | 2     |           |        |             |                 | 1          | 1    | 1         | 3      |               | 1         |                      | 1        |
| <i>topics</i>                   |               |       |           |        |             |                 |            |      |           |        |               |           |                      |          |
| extraction                      |               | 2     |           |        |             | 2               | 4          |      | 1         | 1      | 1             |           |                      | 1        |
| evolution                       | 1             |       |           |        |             |                 | 3          | 1    | 6         |        |               |           |                      | 1        |
| clustering                      | 1             | 2     |           |        |             |                 | 1          | 1    |           | 2      | 1             |           |                      | 3        |
| <i>similar patterns</i>         |               |       |           |        |             |                 |            |      |           |        |               |           |                      |          |
| linguistic patterns             |               | 8     |           |        | 1           | 4               | 1          |      | 5         |        |               | 1         | 2                    | 4        |
| text re-use                     |               | 6     |           |        | 2           | 4               |            |      |           | 3      |               |           |                      | 3        |
| text edition comparison         | 1             | 5     | 2         |        | 4           | 4               |            |      | 1         |        |               | 2         | 2                    | 1        |
| <i>text of interest</i>         |               |       |           |        |             |                 |            |      |           |        |               |           |                      |          |
| interpretation                  |               | 2     | 1         | 1      |             |                 |            |      | 1         |        |               |           |                      |          |
| sound                           |               | 3     |           | 3      | 2           | 3               |            |      |           |        |               |           |                      |          |
| story flow                      | 1             | 1     |           |        |             |                 |            |      | 2         | 1      | 1             |           |                      | 2        |
| miscellaneous                   | 1             | 1     |           | 1      |             | 1               | 1          |      |           | 1      | 1             |           | 1                    | 1        |
| <i>corpus analysis</i>          |               |       |           |        |             |                 |            |      |           |        |               |           |                      |          |
| word statistics & relationships |               | 3     |           |        |             | 2               | 8          |      | 4         | 1      |               |           | 1                    | 2        |
| text similarity                 |               |       |           |        |             | 2               |            |      | 5         |        |               |           |                      |          |
| exploration                     | 2             | 1     | 1         |        |             | 2               | 1          | 1    | 2         | 1      |               |           |                      | 3        |
|                                 | 10            | 38    | 4         | 5      | 9           | 26              | 20         | 18   | 18        | 41     | 9             | 5         | 8                    | 25       |

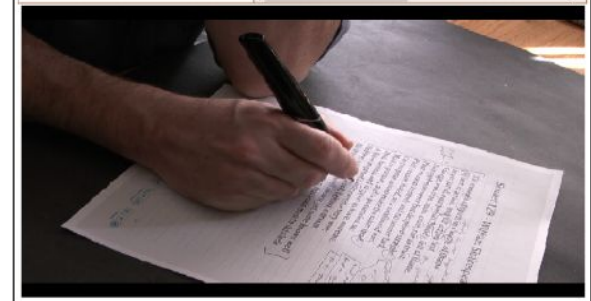
### 3. Tools & Examples

# Tools & Examples: Literature & Linguistics

- Close reading (analysis) of poems traced with a digital pen.
- Phonetic relation tiles, such as consonance, are interactive.
- Different stylistic device are marked with colors and edges.

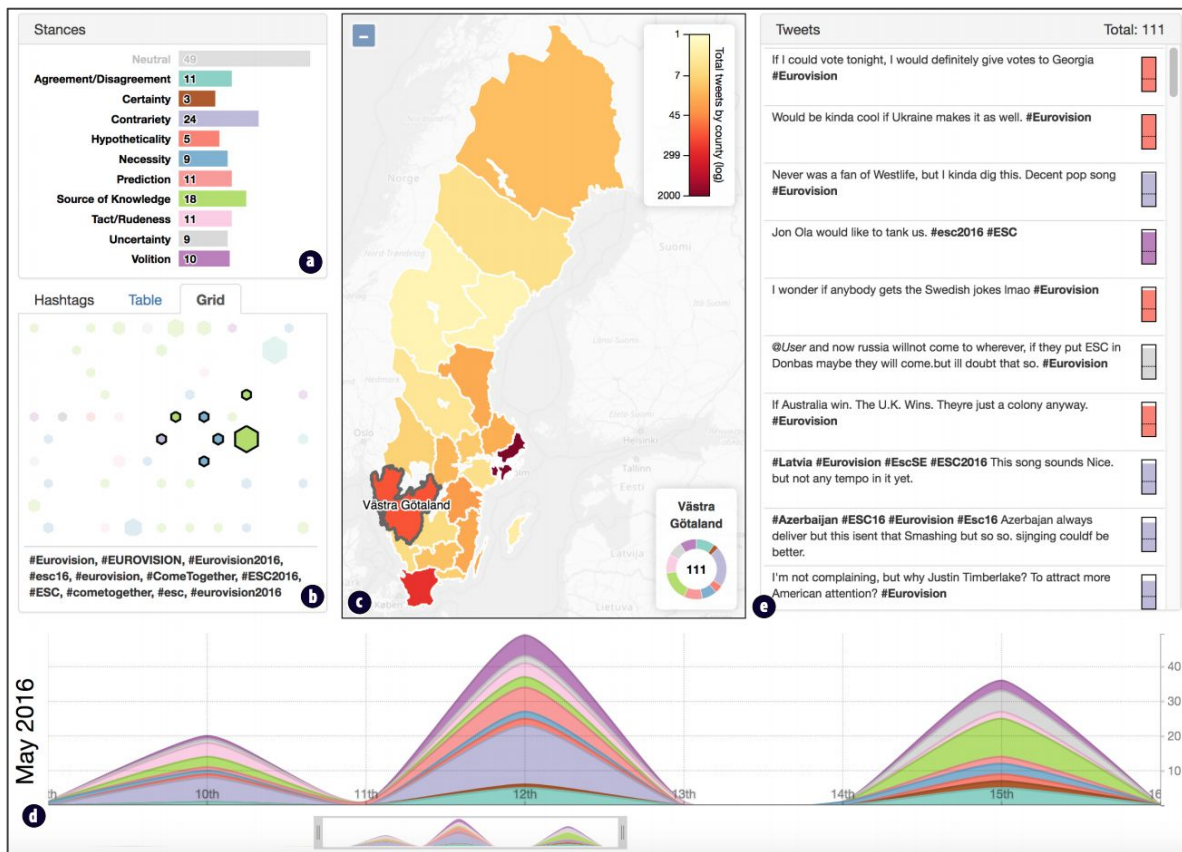
| Consonance   | Perfect rhyme   | Synonyms   | bait   |
|--|---|--|--|
| <p>Sonnet 129</p> <p>William Shakespeare</p> <p>Th' expense of spirit in a waste of shame<br/>           Is lust in action, and till action, lust<br/>           Is perjured, murderous, bloody, full of blame,<br/>           Savage, extreme, rude, cruel, not to trust,<br/>           Enjoys no sinner but despised straight;<br/>           Past reason hunted, and no sooner had,<br/>           Past reason hated, as a swallowed bait,<br/>           On purpose led to make the taker mad;<br/>           Mad in pursuit, and in possession so;<br/>           Had, having, and in quest to have, extreme;<br/>           A bliss in proof, and proved, a very vice;<br/>           Before, a joy proposed; behind, a dream.<br/>           All this the world well knows; yet none knows well<br/>           To shun the heaven that leads men to this hell.</p> | <p>Sonnet 129</p> <p>William Shakespeare</p> <p>Th' expense of spirit in a waste of <b>shame</b><br/>           Is <b>lust</b> in action, and till action, <b>lust</b><br/>           Is perjured, murderous, bloody, full of <b>blame</b>,<br/>           Savage, <b>extreme</b>, rude, cruel, not to <b>trust</b>,<br/>           Enjoys <b>no</b> sinner but despised <b>straight</b>;<br/>           Past reason <b>hunted</b>, and <b>no</b> sooner <b>had</b>,<br/>           Past reason <b>hated</b>, as a swallowed <b>bait</b>,<br/>           On purpose led to make the taker <b>mad</b>;<br/>           Mad in pursuit, and in possession <b>so</b>;<br/>           Had, having, and in quest to have, <b>extreme</b>;<br/>           A <b>bliss</b> in proof, and proved, a very <b>vice</b>;<br/>           Before, a joy proposed; behind, a <b>dream</b>.<br/>           All <b>this</b> the world well knows; yet <b>none</b> knows well<br/>           To <b>shun</b> the heaven that leads men to <b>this</b> hell.</p> | <p>Sonnet 129</p> <p>William Shakespeare</p> <p>Th' expense of spirit in a waste of shame<br/>           Is lust in action, and till action, lust<br/>           Is perjured, murderous, bloody, full of blame,<br/>           Savage, extreme, rude, cruel, not to trust,<br/>           Enjoys no sinner but despised straight;<br/>           Past reason hunted, and no sooner had,<br/>           Past reason hated, as a swallowed bait,<br/>           On purpose led to make the taker mad;<br/>           Mad in pursuit, and in possession so;<br/>           Had, having, and in quest to have, extreme;<br/>           A bliss in proof, and proved, a very vice;<br/>           Before, a joy proposed; behind, a dream.<br/>           All this the world well knows; yet none knows well<br/>           To shun the heaven that leads men to this hell.</p> | <p><b>bait</b></p> <p><b>Origin</b><br/>           Middle English, from Old Norse [beta,] skin to Old English [þetan] to bait, [ðtan] to</p> <p><b>Pronunciations</b><br/>           (BEYL.T)</p> <p><b>Noun</b></p> <p><b>First recorded use</b><br/>           14th century</p> <p><b>Senses</b></p> <p><b>Sense 1</b><br/>           something (as food) used in luring especially to a hook or trap<br/>           a poisonous material placed where it will be eaten by harmful or objectionable pest</p> <p><b>Sense 2</b><br/>           [ture][temptation]</p> <p><b>Verb</b></p> <p><b>First recorded use</b><br/>           13th century</p> <p><b>Senses</b></p> <p><b>Sense 1</b><br/>           to persecute or exasperate with unjust, malicious, or persistent attacks<br/>           [tease]</p> <p><b>Sense 2</b><br/>           to harness (as a chained animal) with does usually for sport</p> |

| Sonnet 129   | Alliteration  |
|--|---|
| <p>Sonnet 129</p> <p>William Shakespeare</p> <p>Th' expense of spirit in a waste of shame<br/>           Is lust in action, and till action, lust<br/>           Is perjured, murderous, bloody, full of blame,<br/>           Savage, extreme, rude, cruel, not to trust,<br/>           Enjoys no sinner but despised straight;<br/>           Past reason hunted, and no sooner had,<br/>           Past reason hated, as a swallowed bait,<br/>           On purpose led to make the taker mad;<br/>           Mad in pursuit, and in possession so;<br/>           Had, having, and in quest to have, extreme;<br/>           A bliss in proof, and proved, a very vice;<br/>           Before, a joy proposed; behind, a dream.<br/>           All this the world well knows; yet none knows well<br/>           To shun the heaven that leads men to this hell.</p> | <p>Sonnet 129</p> <p>William Shakespeare</p> <p>Th' <b>expense</b> of <b>spirit</b> in a <b>waste</b> of <b>shame</b><br/>           Is <b>lust</b> in action, and till action, <b>lust</b><br/>           Is <b>perjured</b>, <b>murderous</b>, <b>bloody</b>, full of <b>blame</b>,<br/> <b>Savage</b>, <b>extreme</b>, <b>rude</b>, <b>cruel</b>, not to <b>trust</b>,<br/>           Enjoys <b>no</b> sinner but <b>despised</b> <b>straight</b>;<br/> <b>Past</b> reason <b>hunted</b>, and <b>no</b> sooner <b>had</b>,<br/> <b>Past</b> reason <b>hated</b>, as a <b>swallowed</b> <b>bait</b>,<br/>           On <b>purpose</b> led to make the taker <b>mad</b>;<br/>           Mad in <b>pursuit</b>, and in <b>possession</b> <b>so</b>;<br/>           Had, having, and in <b>quest</b> to have, <b>extreme</b>;<br/>           A <b>bliss</b> in <b>proof</b>, and <b>proved</b>, a very <b>vice</b>;<br/>           Before, a <b>joy</b> <b>proposed</b>; behind, a <b>dream</b>;<br/>           On <b>purpose</b> led to make the taker <b>mad</b>;<br/>           Mad in <b>pursuit</b>, and in <b>possession</b> <b>so</b>;<br/>           Had, having, and in <b>quest</b> to have, <b>extreme</b>;<br/>           A <b>bliss</b> in <b>proof</b>, and <b>proved</b>, a very <b>vice</b>;<br/>           Before, a <b>joy</b> <b>proposed</b>; behind, a <b>dream</b>.<br/>           All <b>this</b> the world well knows; yet <b>none</b> knows well<br/>           To <b>shun</b> the heaven that leads men to <b>this</b> hell.</p> |



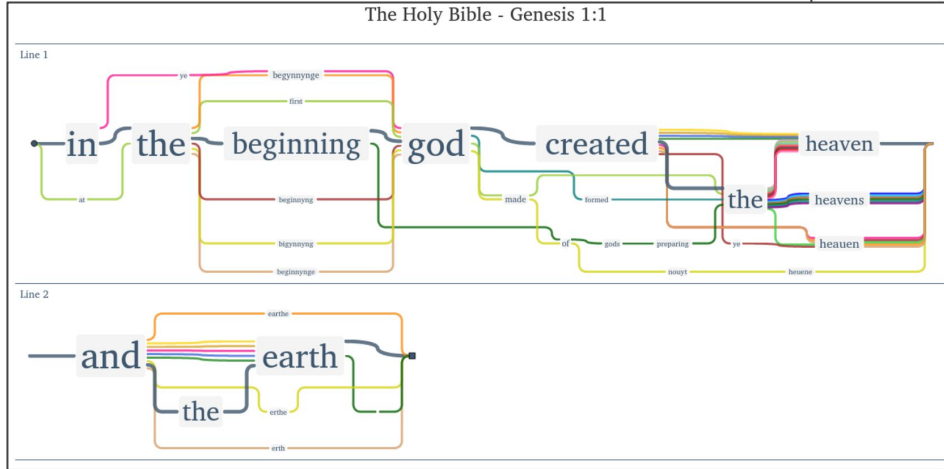


# Tools & Examples: Communication & Political Science

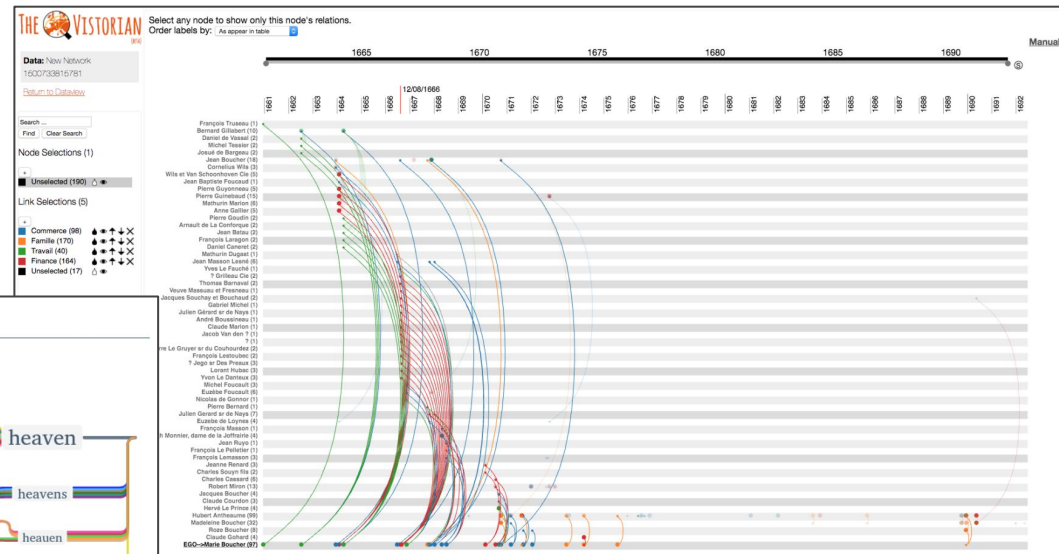


- Visualization for the interactive exploration of stance in social media.
- Coordinated multiple views including user-defined topics, content similarity and dissimilarity, and geographical and temporal distribution.
- Investigation multidimensional patterns and trends in stance-taking related to cultural events.

# Tools & Examples: History



Franzini, G., Terras, M., Mahony, S., Jänicke, S., Scheuermann, G., & Geßner, A. (2015). TRAViz: A visualization for Variant Graphs. *Digital Scholarship in the Humanities*, 30.



VS Molinero, B Bach, C Plaisant, N Dufournaud, JD Fekete. Understanding the Use of The Vistorian: Complementing Logs with Context Mini-Questionnaires (2017).

- Tools to allow digital humanists to analyze complex geolocated and temporal networks of agents, works and institutions.
- Visualizing variation between editions of historical documents and manuscripts.

# Austrian Media Corpus



CC-BY 4.0, Sandra Lehecka, <https://amc.acdh.oew.ac.at/publikationen/>

- A corpus (plural: corpora) is a collection of written or spoken text.
- Corpus-based research is a key element of almost all language studies.
- This corpus can be categorized under contemporary media corpora.

# Austrian Media Corpus



- Created by ÖAW & APA and covers the “entire” Austrian written media of the past 20 years.
- Contains 40 million documents and more than 10 billion tokens.

# Austrian Media Corpus

## Coverage: Contained Media

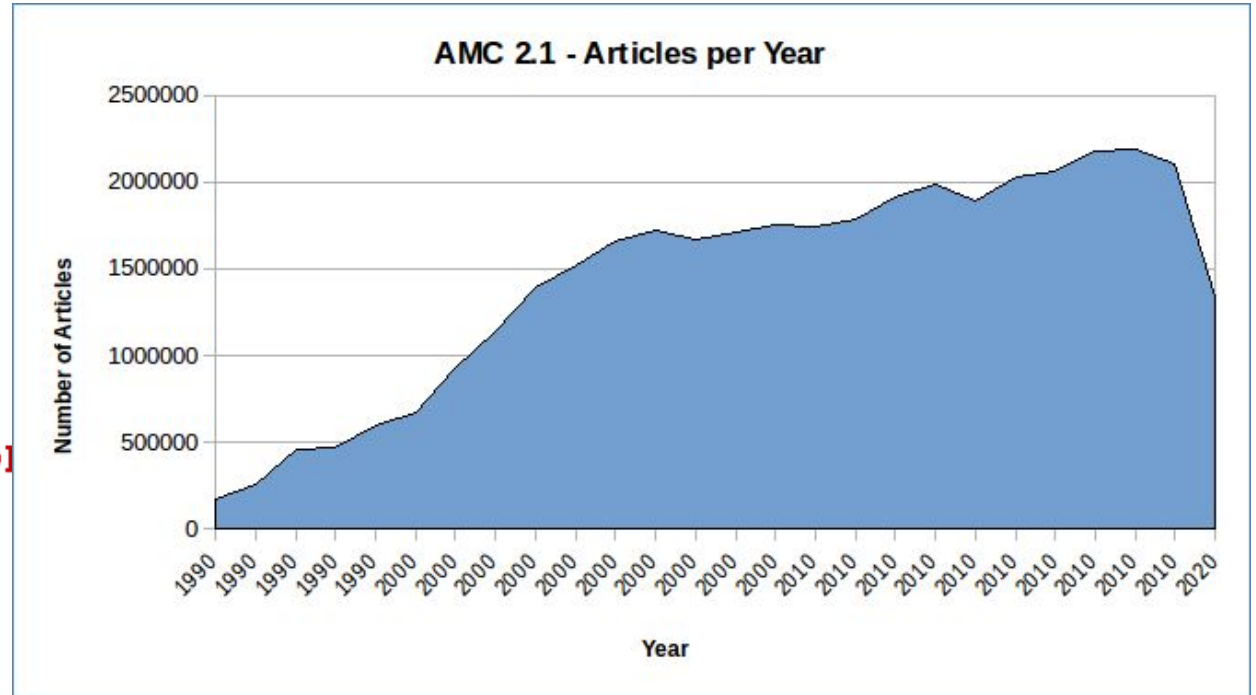


|  |   |  |
|--|---|--|
| <u>National Newspapers:</u><br>Der Standard<br>Die Presse<br>Heute<br>Kurier<br>Kronen Zeitung<br>Kleine Zeitung<br>Medianet<br>Neues Volksblatt<br>Österreich<br>Wiener Zeitung<br>Wirtschaftsblatt | <u>Regional Newspapers:</u><br>Kärntner Tageszeitung<br>Neue Vorarlberger Tageszeitung<br>Niederösterreichische<br>Nachrichten<br>Oberösterreichische<br>Nachrichten<br>Salzburger Nachrichten<br>Salzburger Volkszeitung<br>Tiroler Tageszeitung<br>Vorarlberger Nachrichten | <u>Weekly Papers:</u><br>Falter<br>Die Furche<br>Sportzeitung<br>Burgenländische Volkszeitung<br>Oberländer Rundschau<br>Salzburger Woche<br>Bauernzeitung<br>Kärntner Wirtschaft<br>Solidarität                         |
| <u>Periodicals:</u><br>Arbeit & Wirtschaft<br>Augustin<br>Die Wirtschaft   | <u>Weekly Magazines:</u><br>Format<br>News<br>Profil<br>TV-Media<br>woman<br>E-media<br>Der Grazer<br>Horizont  | <u>Monthly Magazines:</u><br>Academia<br>Arbeit und Wirtschaft<br>Datum<br>Die Wirtschaft<br>Echo<br>Gewinn<br>Trend<br>Wiener<br>Wienerin<br>Neuer Kärntner Monat<br>Steirer Monat<br>Industriemagazin<br>Der Konsument |

# Austrian Media Corpus

Coverage: Temporal

**amc**  
austrian media corpus



# Austrian Media Corpus

## Underlying Data

- **Metadata about article**
  - Media name
  - Region
  - Date
  - Ressort
  - Page Number
  - (Author)
- **Annotations on the text**
  - Lemma
  - Part-of-Speech tag(s!)
  - Morphological info
  - Named Entity (PERS, LOC, ORG)
  - Dependency Parse-Tree

# Underlying Data

schenken /VVFIN/schenken die /ART/die Menschen /NN/Mensch den /ART/die im /APPRART/in „ /\$(/' Haus /NN/Haus der /ART/die Gesundheit /NN/Gesundheit “ /\$(/' ansässigen /ADJA/ansässig

| Unit Sizes of the Corpus |                |
|--------------------------|----------------|
| Unit                     | Count          |
| Documents                | 42.240.787     |
| Paragraphs               | 317.301.865    |
| Sentences                | 1.154.897.537  |
| Words                    | 7.859.278.624  |
| Tokens                   | 10.500.118.851 |

Total unit sizes and an example document from the Austrian Media Corpus.

Sieben Ärzte auf einen Streich in Hof bei Salzburg  
Von Kopf bis Fuß gut aufgehoben sind die Menschen seit sieben Jahren im „ Haus der Gesundheit “ - Ärztezentrum in Hof .  
Vom Allgemein-Mediziner bis zum Facharzt findet man hier alles unter einem Dach . Mit der Apotheke wurden im „ Haus der Gesundheit “ 30 Arbeitsplätze geschaffen . Kurze Wege , prompte und zeitnahe Versorgung garantieren beste Betreuung der Patienten .

Das Team um Apothekenleiter und Initiator Dr. Wolfgang Beindl freut sich über den großen Zuspruch aus der Region . Von Hof bis Bad Ischl und aus der Stadt Salzburg schenken die Menschen den im „ Haus der Gesundheit “ ansässigen Medizinern ihr Vertrauen und nutzen das Angebot der echten Nahversorgung für die Gesundheit .

Dr. Paul Windischbauer , Wahlarzt für Hals- , Nasen- , - und Ohrenheilkunde , eröffnete vor einer Woche seine technisch modernst ausgestattete Ordination im Haus . Ohne Wartezeit nimmt er sich für Anamnese und Diagnostik der Beschwerden ausreichend Zeit . „ In angenehmer Atmosphäre , stressfrei ohne Zeitdruck , führe ich jede Untersuchung schonend durch “ , sagt er . Auch Kindern kommen die einfühlsamen und stressfreien Untersuchungen zugute .

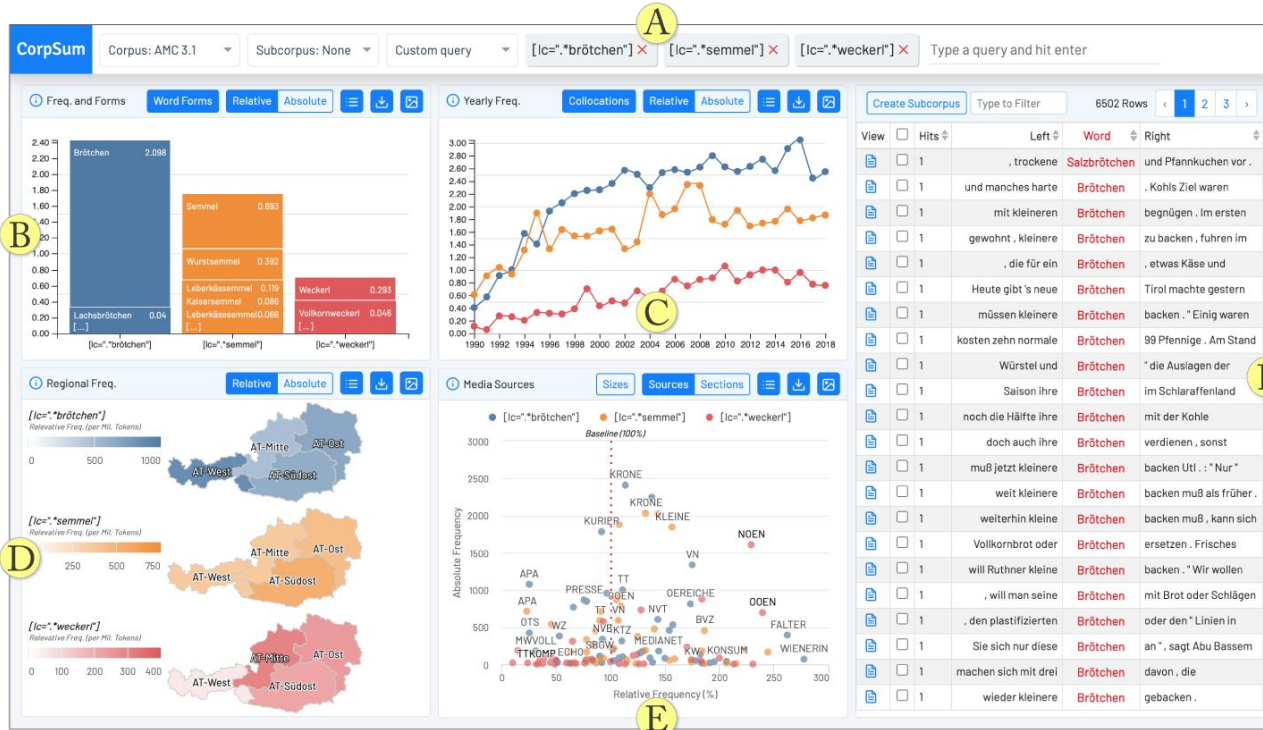
Dr. Paul Windischbauer eröffnete seine Ordination im „ Haus der Gesundheit “ Nach 20 Jahren an den SALK und als Oberarzt der HNO- Univ. Klinik Salzburg hat Dr. Paul Windischbauer vor zwei Jahren seine Wahlarzt-Heimat in Hof gefunden . Nun hat er seine Praxis direkt im Ärztezentrum . Die Schwerpunkte des auch in der EMCO-Klinik und am Diakonissen-Krankenhaus praktizierenden Arztes sind die allgemeine Rhinologie , die ästhetische und funktionelle Nasenchirurgie bei Atemproblemen , Nasennebenhöhlenchirurgie , die Tumorchirurgie etc. Im Hörtraum führt er Schallmessungen und Trommelfellmessungen durch .

Der menschliche Umgang mit Patientinnen und Patienten ist ein zentrales Anliegen . Das gesamte Spektrum von Behandlungen - bis auf Haut- und Zahnbehandlungen - ist im Haus der Gesundheit zu bekommen . Hoyer ab . „ Patienten kommen so ohne Umwege zu uns . Termine sind frei gehalten .

Zusätzlich zum umfassenden Ärzte-Angebot bemühen sich die Mitarbeiterinnen und Mitarbeiter um eine optimale Ernährungsbereitstellung . Ernährungsberaterin Maria Anna Benedikt um die Ernährung der Patienten . Zudem finden zahlreiche Vorträge und Seminare statt . Auch die Öffnungszeiten sind besonders kundenfreundlich . Besonders stolz ist er über das innovative Netzwerk der Haus der Gesundheit . Unverträglichkeiten , helfen uns die zahlreichen Daten zu analysieren . „ , sagt der klinische Pharmazeut .

Das fächerübergreifende Zusammenarbeiten in der Haus der Gesundheit . Gibt es in Krankenhäusern oft lange Wartezeiten , kann man hier im Haus der Gesundheit . Und die Apotheke ? Sie ist Drehscheibe für einen un

|                 |  |
|-----------------|--|
| Token number    | 10494826104  |
| Document number | 42224075   |
| doc.id          | SBGW_20181122030818226100021                       |
| doc.datum_full  | 2018-11-22T00:00:00Z                               |
| doc.datum       | 2018-11-22   |
| doc.bibl        | Salzburger Woche 47 vom 2018-11-22 s. 12           |
| doc.docsrc      | SBGW   |
| doc.docsrc_name | Salzburger Woche                                   |
| doc.region      | amitte   |
| doc.tokens      | 435  |
| doc.ressort2    | allgemein  |
| doc.mutation    | Flachgauer Nachrichten                             |
| doc.keys        | SE_12 DA_20181122 MO_201811 JA_2018 RS_ALL DB_SBGW |
| doc.mediatype   | print  |



**An enabling visual analysis tool-design to help the corpus-researchers to conduct their work easier and faster.**

## 4. Live Exercises with Voyant



## 5. Discussions & Challenges

From the computational and quantitative analysis of “low-hanging fruits” - to a completely new AI ball game..



