

# AnnotateVis: Combining Traditional Close Reading with Visual Text Analysis

M. Faisal Cheema, Stefan Jänicke, and Gerik Scheuermann, *Member, IEEE*

**Abstract**—Close reading is an important research method to analyze and interpret text in traditional humanities research. In the last decades, a huge amount of historical texts have been made available due to large scale digitization projects and facilitated distant reading methods. In the last decade, a number of visualization techniques were developed that support distant reading tasks as they aid to explore large text collections. But, less work has been done to improve the close reading capabilities of humanities scholars, and the majority of existing techniques are either advanced text viewers highlighting textual features, or they are specific to particular research tasks. We propose the close reading tool AnnotateVis that supports the annotation of text fragments with various media types (audio, video, images, etc.). In addition, a user can perform various text analysis methods for a desired text snippet, and the results can be annotated in the form of an interactive visualization. To illustrate the utility of AnnotateVis, we provide two close reading scenarios.

**Index Terms**—Close Reading, Annotation tool, Distant Reading

---

## 1 INTRODUCTION

Textual analysis is a fundamental practice in different disciplines of humanities research to analyze historical texts and narratives. Since the early twentieth century, *Close Reading* has been one of the significant research methods to analyze text in traditional humanities research - mainly literary criticism [15]. Close reading has a long tradition. It goes back to Aristotle who close read the works of Plato [21]. A definition of close reading is given by Nancy Boyles [6]: *Essentially, close reading means reading to uncover layers of meaning that lead to deep comprehension*. Although the close reading is steered by the scholars intellectual thought process, there are several techniques, methods and theories to approach it [7]. Close reading constitutes of a range of methodologies e.g.: multiple readings of the text, placing the text in its historical and socio-historical context, looking for patterns (e.g., repetitions, contradictions and similarities), finding inherent topics and occurring relationships, etc. Traditionally, close reading is done on paper. Annotating the text is considered as a strong facilitating method in a close reading process [27].

At the end of the twentieth century, there were great advancements in the field of computing. There was keen interest from humanities scholars and computer scientists to digitize historical texts for preservation purposes and to support collaborative work on those digital editions. This cooperation established the digital humanities field - an intersection of computer science and humanities. A huge amount of historical texts were made available due to large scale digitization projects and provided the scholars working in the digital humanities with a number of fascinating prospects. With huge text sources and corpora available, the digital humanities revolutionized textual analysis methodologies in comparison to classical methods. Instead of focusing on a single piece of text, there was a growing desire to explore this sea of historical texts and discover trends and outliers. A major boost to this thought was given by Moretti, when he pioneered the term *Distant Reading* [24]. The idea of distant reading was put forth by Moretti from the fact that there are just too many texts to be seriously read and studied. In our survey [17] we define distant reading as in contrast to close reading as follows: *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)*. In his book, Moretti showed different illustrations in form of graphs, maps and trees - thereby advocating the use of visualizations for distant reading tasks. After Moretti published his idea, a number of visualization techniques were developed in the context of digital humanities, as summarized in our survey [17]. We were also able to show that the method of distant reading was gradually embraced by the digital humanities community and the majority of developed visualization techniques aided the exploration of large text collections. We further observed that less work has been done to improve the close reading capabilities of humanities scholars. The majority of existing close reading techniques acted merely as advanced text viewers highlighting textual features. On the other hand, there are close reading tools, e.g., TRAViz [18], Poem Viewer [1] or Poemage [22], that are pertinent to a particular research task.

Even after rapid developments in distant reading, our survey showed that close reading still remains a prior task in humanities research. Even in digital humanities projects that aimed to develop a distant reading technique in the first place, it was very important for the involved humanities scholars to actually see the underlying text sources, which is a mandatory functionality in order to verify distant reading hypotheses and to build trust in the novel research approach. That visualization can play important role to support the close reading of texts has already been shown by [13], who proposes so called word scale visualizations and annotates individual words computationally with statistical or geographical information. We propose *AnnotateVis*, an enhanced close reading tool that rather supports the traditional close reading workflow of humanities scholars digitally, i.e., by being able to annotate and highlight text fragments using various methods. In addition, we provide various visual analysis methods for desired text snippets in the close reading environment. The design of the tool is inspired by mind maps, putting the text in the center of the screen and arranging user annotations strategically around it to support the scholars individual intellectual thought process.

In the section 2, we illustrate the traditional close reading approach using annotations and observe existing tools that support annotating digital texts. In addition, we discuss how the close reading experience can be enhanced by providing various distant reading views on the text. In section 3, we discuss the design and features of AnnotateVis, and in section 4, we provide two usage scenarios that illustrate the benefit of our approach.

## 2 RELATED WORK

Close reading is a process where the humanities scholar carefully reads and interprets a text. In order to facilitate the understanding of a text fragment, annotating the text to highlight as well as to record

- 
- M. Faisal Cheema is with Leipzig University. E-mail: [faisal@informatik.uni-leipzig.de](mailto:faisal@informatik.uni-leipzig.de).
  - Stefan Jänicke is with Leipzig University. E-mail: [stjaenicke@informatik.uni-leipzig.de](mailto:stjaenicke@informatik.uni-leipzig.de).
  - Gerik Scheuermann is with Leipzig University. E-mail: [scheuermann@informatik.uni-leipzig.de](mailto:scheuermann@informatik.uni-leipzig.de).





Fig. 3. AnnotateVis screenshot: User close reading text snippet from *David Copperfield*

of common mind maps, the central node does not depict any meaning. Insofar, the mind map usage in a close reading process differs as the text in question conveys a lot of information. The prominent features of AnnotateVis are as follows:

**Integrated environment** AnnotateVis provides an integrated environment where the scholar can annotate and perform text analyses in multifarious ways. In addition to annotating text fragments as also provided by other web annotation tools, the scholar can additionally create annotations in the form of media such as audio, video and pictures, which support the scholars close reading process. Notably, the scholar can perform a number of text analyses using distant reading tools such as tag clouds, charts, maps, timelines etc. and place these annotations as interactive visualizations. All annotations are placed in the margins alongside the text.

**Free layout** Following the design strategy of mind maps, AnnotateVis provides a free layout for the placement of annotations. The scholar can decide the best suited position for an annotation according to its importance and its proximity to the text. As opposed to the traditional close reading, which results in a final setting of annotations,

the margins alongside the digital text can be flexibly annotated, and the layout can be changed interactively.

**Distant reading techniques** AnnotateVis offers a multitude of distant reading tools. Thereby, the scholar can select an appropriate visualization technique in accordance to the observed text fragment and the research question at hand. This is an important feature as closed reading is a customizable process.

**Annotation scope** Another important feature of AnnotateVis is the ability of performing text analyses at various text hierarchy levels. Depending on the research question as well as the selected text analysis method, the scholar can select a text fragment to be examined, e.g., the current page, a paragraph, a book chapter, or the entire text.

**Workflow integration** As mentioned before, AnnotateVis is designed the way that it is easy for humanities scholars to integrate a close reading into their existing research workflows. The scholar can save the results and reload them for further analysis. This also provides the opportunity to share the individual close reading with other scholars, and supports the collaborative work essential for (digital) humanities scholars.

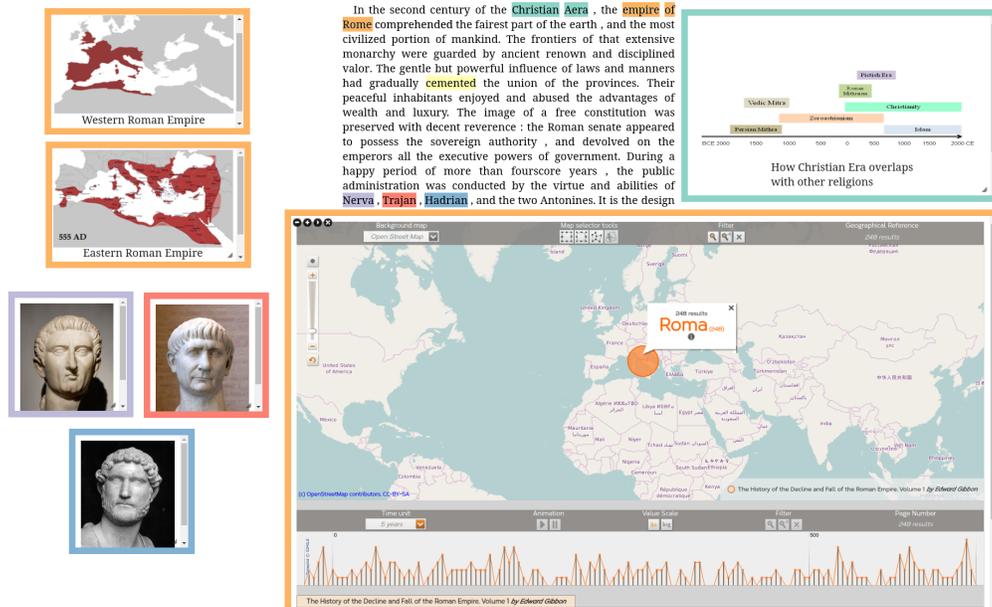


Fig. 4. AnnotateVis screenshot: User close reading text snippet from *History of the decline and fall of the Roman Empire*

#### 4 USAGE SCENARIOS

In this section, we explain the capabilities of AnnotateVis with two usage scenarios. In Fig. 2, a text excerpt from Charles Dickens *David Copperfield* is close read by a humanities scholar. We see how the scholar annotated observations, assertions and assumptions in the close reading process. Fig. 3 shows the same portion of text being close read in AnnotateVis. The user selected text fragments annotated with additional information to the observed subject. The color of the annotated text and the borders of the annotation boxes are used as the visual identifiers for the user. Below, we list the users annotations of text fragments, which are shown in Fig. 3.

- In the top left we see a video annotation (light green), which embeds a YouTube source containing an audio recording of chapter 2.
- The user adds a map showing different places where David lived in his *childhood*. The annotation border and the term childhood in the text are marked in blue.
- In the traditional close reading scenario explained in Fig. 2, the scholar made the assumption if *memory* equals *observation*. In AnnotateVis, the user applied TagPies [16] to further investigate this assumption. As shown in the top right corner of Fig. 3 in purple, the user created three different TagPies for the terms memory, observe and remember. Their co-occurrences are compared on different text hierarchy levels: the currently analyzed chapter 2, a subset of chapters, and the entire book.
- The user selected a paragraph of the text, and annotated his interpretation in the form of digital text annotation (grey).
- To illustrate the phrase *here is our pew at a church*, the user annotates an image (marked in green) taken from a printed edition.
- On the right bottom left, the user annotated the term *Peggotty* in cyan color. Firstly, the user wrote his observations about Peggotty's character. Secondly, he put an image of Peggotty found in the web. Thirdly, the user added an interactive annotation in form of concordance plot to see how Peggotty was mentioned in the concerning chapter 2.

In the second scenario, the user close read a text fragment of Edward Gibbon's *History of the decline and fall of the Roman Empire*. A

screenshot of the close reading process in AnnotateVis is shown in Fig. 4. The user analyzed the text's concept of the *Roman Empire*. At first, he added two maps of the Western and Eastern Roman Empire of envisage its temporal change of geographical dimension. Furthermore, the user analyzed the usage of the placename *Rome* in the book using GeoTemCo [19]. The time graph shows how the word Rome was used throughout the book. Usually, the interactive annotations are placed in a minimized view next to the text, but the user can also enlarge a view (map) for further analysis. Another annotation in light green (top right) shows a timeline that illustrates how the *Christian era* coincides with other religions. Finally, we see that user annotated images of Emperors *Nerva*, *Trajan* and *Hadrian*.

#### 5 FEEDBACK

We discussed the AnnotateVis design and functionality with collaborating humanities scholars. They mentioned that the proposed design resembles the traditional close reading on paper, and that AnnotateVis will help humanities scholars to quickly adapt and integrate this digital close reading methodology into their existing workflows. This could help lower the fear of scholars in classical studies who are reluctant to use computers in their traditional workflows. We also presented the proposed design of AnnotateVis at the Digital Humanities conference 2016 [10] and discussed the design with various humanities scholars, who appreciated our novel close reading methodology, stating that there is a need for a close reading tool that supports distant reading visualizations. We also got positive feedback concerning the placement strategy of annotations. One of the scholars remarked that it would benefit them to group the annotations as they desire, instead of a tool forcing them to use a predefined layout.

#### 6 FUTURE WORK

At the moment, AnnotateVis provides basic text analysis tasks supported by distant reading visualizations. In the future, we intend adding information extraction techniques such as Named Entity recognition, Relationship Extraction or Terminology extraction, so that the text analysis base can be broadened and more distant reading visualizations can be added to the existing toolbox to further facilitate the humanities scholars close reading experience. At the moment, the placement of annotations is user dependent. We plan to improve the layout strategy by suggesting an optimal position for an annotation. In addition, we intend to perform a user study to evaluate the utility of AnnotateVis to support scholars in close reading, and to spot necessary improvements.

## REFERENCES

- [1] A. Abdul-Rahman, J. Lein, K. Coles, E. Maguire, M. Meyer, M. Wynne, C. R. Johnson, A. Trefethen, and M. Chen. Rule-based Visual Mappings—with a Case Study on Poetry Visualization. In *Computer Graphics Forum*, vol. 32, pp. 381–390. Wiley Online Library, 2013.
- [2] M. Agrawala and M. Shilman. DIZI: A Digital Ink Zooming Interface for Document Annotation. In *IFIP Conference on Human-Computer Interaction*, pp. 69–79. Springer, 2005.
- [3] B. Alex, C. Grover, J. Oberlander, K. Zhou, and U. Hinrichs. Palimpsest: Improving Assisted Curation of Loco-specific Literature. In *Proceedings of Digital Humanities 2015, 2-3 July, Sydney, Australia*, 2015.
- [4] D. Barger and T. Moscovich. Reflowing Digital Ink Annotations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 385–393. ACM, 2003.
- [5] M. Bonn and J. McGlone. New Feature: Article Annotation with Hypothes.is. *Journal of Electronic Publishing*, 17(2), 2014.
- [6] N. Boyles. Closing in on Close Reading. *Educational Leadership*, 70(4):36–41, 2012.
- [7] B. Brummett. *Techniques of Close Reading*. Sage, 2009.
- [8] J. W. Budd. Mind Maps as Classroom Exercises. *The Journal of Economic Education*, 35(1):35–46, 2004.
- [9] T. Buzan and B. Buzan. *The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential*. Plume, 1996.
- [10] M. Cheema, S. Jänicke, and G. Scheuermann. Enhancing Close Reading. In *Digital Humanities*, pp. 758–761, 2016.
- [11] R. Enslin. Mindmap: Coffee Project, 2010. [Online; accessed August 6, 2016].
- [12] Z. Geng, T. Cheesman, R. S. Laramée, K. Flanagan, and S. Thiel. ShakerVis: Visual Analysis of Segment Variation of German Translations of Shakespeare's Othello. *Information Visualization*, 14(4):273–288, 2015.
- [13] P. Goffin, W. Willett, J.-D. Fekete, and P. Isenberg. Exploring the Placement and Design of Word-Scale Visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2291–2300, Dec 2014.
- [14] M. Grassi, C. Morbidoni, M. Nucci, S. Fonda, and F. Di Donato. Pundit: Creating, Exploring and Consuming Semantic Annotations. In *SDA*, pp. 65–72. Citeseer, 2013.
- [15] J. Hawthorn. *A Glossary of Contemporary Literary Theory*. Arnold, 1998.
- [16] S. Jänicke, J. Blumenstein, M. Rücker, D. Zeckzer, and G. Scheuermann. Visualizing the Results of Search Queries on Ancient Text Corpora with Tag Pies. *Digital Humanities Quarterly*, 2016.
- [17] S. Jänicke, G. Franzini, M. Cheema, and G. Scheuermann. Visual Text Analysis in Digital Humanities. In *Computer Graphics Forum*. Wiley Online Library, 2016.
- [18] S. Jänicke, A. Geßner, G. Franzini, M. Terras, S. Mahony, and G. Scheuermann. TRAViz: A Visualization for Variant Graphs. *Digital Scholarship in the Humanities*, 30(suppl 1):i83–i99, 2015.
- [19] S. Jänicke, C. Heine, and G. Scheuermann. GeoTemCo: Comparative Visualization of Geospatial-Temporal Data with Clutter Removal Based on Dynamic Delaunay Triangulations. In *Computer Vision, Imaging and Computer Graphics. Theory and Application*, pp. 160–175. Springer, 2013.
- [20] A. Kehoe and M. Gee. eMargin: A Collaborative Textual Annotation Tool. *Ariadne*, (71), 2013.
- [21] M. M. McCabe. *Platonic Conversations*. Oxford University Press, USA, 2015.
- [22] N. McCurdy, J. Lein, K. Coles, and M. Meyer. Poemage: Visualizing the Sonic Topology of a Poem. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1):439–448, Jan 2016.
- [23] MindTools.com. Mindmap: Time management, 2008. [Online; accessed August 6, 2016].
- [24] F. Moretti. *Graphs, Maps, Trees: Abstract Models for a Literary History*. Verso, 2005.
- [25] A. Muralidharan. A Visual Interface for Exploring Language Use in Slave Narratives. *Digital Humanities 2011*, p. 338, 2011.
- [26] J. C. Nesbit and O. O. Adesope. Learning With Concept and Knowledge Maps: A Meta-Analysis. *Review of educational research*, 76(3):413–448, 2006.
- [27] C. Porter-O'Donnell. Beyond the Yellow Highlighter: Teaching Annotation Skills to Improve Reading Comprehension. *English Journal*, pp. 82–89, 2004.
- [28] B. N. Schilit, G. Golovchinsky, and M. N. Price. Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 249–256. ACM Press/Addison-Wesley Publishing Co., 1998.
- [29] S. Sinclair, G. Rockwell, et al. Voyant Tools. *Online: http://voyant-tools.org*, 2012.
- [30] D. Yoon, N. Chen, and F. Guimbretière. TextTearing: Expanding Whitespace for Digital Ink Annotation. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pp. 107–112. ACM, 2013.
- [31] S. Zyto, D. Karger, M. Ackerman, and S. Mahajan. Successful Classroom Deployment of a Social Document Annotation System. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1883–1892. ACM, 2012.