

AnnoDiver: Applying Visual Analytics on Social Annotations to Facilitate Balanced Research Paper Discourse

Jasmine Shih*
Stanford University

Nick Feffer†
Stanford University

Miroslav Suzara‡
Stanford University

Kevin Wang§
Stanford University

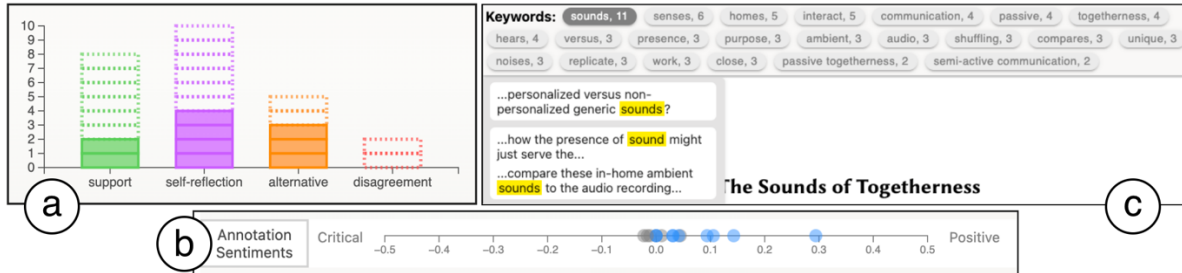


Figure 1: Three social annotation visualizations in AnnoDiver: (a) category bar chart, (b) sentiment axis, and (c) key term frequency list.

ABSTRACT

Prior work has shown that it can be difficult to navigate a large volume of social annotations and that establishing diverse discussions remains a challenge, particularly in university classroom settings. To address this, we developed a prototype social annotation tool that displays interactive visualizations of annotation threads. The visualizations included: 1) a bar chart that displays the counts of different categories of comments, 2) a sentiment axis that shows the sentiment distribution, and 3) a key term frequency list that displays the most frequent terms in the comments. In a user study, we found that users who interacted with the visualizations explored more annotation threads and created more comments, although the diversity of categories of comments created did not change. Our results provide design insights and inspire future research directions on the roles of interactive visualizations on social annotation data.

Keywords: Social annotation, collaborative learning, text/document data, information visualization

Index Terms: Human-centered computing—Collaborative and social computing—Collaborative and social computing systems and tools; Human-centered computing—Visualization—Visualization application domains

1 INTRODUCTION

Reading and reflecting on research texts is often an isolated activity, with only limited opportunities to learn from peers in places outside of the text, such as class discussions. Social annotation tools can address this by embedding individual and collective thinking within the text [1][2]. However, prior research has shown that it can be difficult to navigate this abundance of

annotations and that learners were hesitant about expressing disagreement with others [3], limiting the ability of social annotations to serve as a scaffold for more balanced class-wide discussion that is inclusive of diverse viewpoints.

We hypothesized that integrating a series of interactive visualizations of social annotation threads in a research paper annotation tool would aid annotation navigation and lead to a more diverse discussion. The contributions of this work are the design of AnnoDiver, a novel social annotation system focused on research texts with integrated and interactive visualizations, as well as quantitative and qualitative insights gained from our preliminary user study. Given the design probe of AnnoDiver, we came up with two hypotheses that we tested in the user study:

H1: The interactive visualizations of the comments will allow and encourage readers to explore more annotation threads.

H2: Displaying the total counts of comments per category will motivate readers to leave comments in less-populated categories.

2 SYSTEM

AnnoDiver operates as a web interface built on top of [react-pdf-highlighter](#), a React library that provides basic functionality for PDF highlighting and annotating. Users can annotate (i.e. highlight) text on a PDF and respond to others' annotations in a thread. The system prompts users to categorize their annotation and reply comments by type (*support*, *self-reflection*, *alternative*, and *disagreement*, as used in Gao's case study [3]) and presents explorable visualizations of this text data built with [D3.js](#). Upon clicking on an annotation highlight or a visualization element that represents a comment, the user is presented with the annotation thread that the selected comment belongs to in the ThreadView. Additionally, the visualization elements that all represent the same comment in the three visualization views are linked and highlighted in yellow to allow the user to easily locate them.

2.1 Comment Category Bar Chart

The comment category bar chart is an interactive visualization that displays comment counts in each of the four categories on the current paper. Each bar indicates the count of a comment category and is further divided into rectangles that represent distinct comments. To use the visualization for navigation, the user can click on any rectangle to jump to that respective comment. The comments authored by the currently logged-in user are

*e-mail: jyshih@stanford.edu

†e-mail: nfeffer@stanford.edu

‡e-mail: msuzara@stanford.edu

§e-mail: kevwang1@stanford.edu

(All authors contributed equally to this research.)

distinguished by being filled with a solid color and displayed at the bottom of the bar, while the comments created by other users are displayed with a dotted outline. The stacked bar chart design allows users to keep track of their own contributions and see how their contribution patterns compare with the rest of the group.

2.2 Sentiment Axis

The sentiment axis displays the distribution of the sentiments conveyed in the annotation comments on the paper. Each annotation comment is plotted on the axis (ranging from -0.5 to 0.5) as a circle based on its sentiment score, which we obtain by passing the entire comment into the [sentiment Node.js module](#). In addition to seeing the group’s overall sentiments towards the paper, the user can navigate the annotations by clicking on any circle on the axis. Similar to the comment category bar chart design, the user’s own comments are colored in blue, as opposed to the gray that other users’ comments are colored in.

2.3 Key Term Frequency List

The key term list shows a quickly digestible synthesis of peer commentary and also functions as a navigation tool. Each key term can be clicked to open a drop-down window that shows all of the comments where the term appeared. When the user clicks on one of the comments in the drop-down, the ThreadView is updated to show the thread in which that comment appeared. The key terms shown in the list are generated using the [keyword-extractor Node.js module](#) which scans every sentence for one and two-word terms, aggregating a frequency list of every term. We then further process the frequency list with pluralization detection, so that words and their plural forms are treated as the same term.

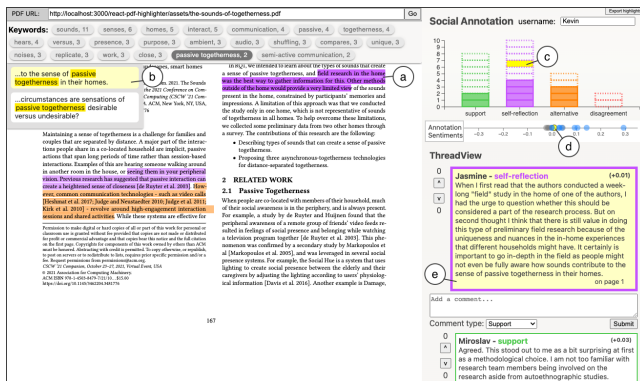


Figure 2: User interface after a comment has been selected: (a) shows the highlight that the view has been scrolled to, (b-d) show the visualization elements that represent the selected comment highlighted in yellow, and (e) shows the selected comment highlighted in the thread.

3 EVALUATION

3.1 Methodology

We conducted a user study with 8 university students who were separated into a control group (n=3), which received a version of AnnoDiver without the visualizations, and an experimental group (n=5) with the visualizations. Our research team selected a four-page paper and seeded 15 annotations as well as 10 replies across all categories. We intentionally left an imbalance that included an abundance of *self-reflection* and *support* annotations and less *alternative* and *disagreement* annotations, in order to prompt the participants toward categories that are typically less engaged with in social annotation contexts. During 45-minute Zoom sessions,

each participant was asked to complete two tasks: 1) create at least three comments, and 2) find three different viewpoints presented by the paper or discussed in the social annotations.

3.2 Findings

3.2.1 Quantitative Differences between Groups

We computed the average number of annotation threads the participants in each group viewed and the average number of comments the participants in each group created. We then compared these metrics between the two groups. For annotation exploration, the control group viewed an average of 7.3 annotations while the experimental group viewed an average of 8.6 annotations. Although not statistically significant, this difference between the two groups suggests that the visualization features might have led the participants to explore more annotation threads. For comment creation, the control group created an average of 4 comments while the experimental group created an average of 5.8 comments. This difference, though not statistically significant, suggests that the visualization features might have led the participants to create more annotations on the paper and respond more to the pre-seeded social annotations.

3.2.2 Visualization Usage Patterns

Among the visualizations, participants most commonly used the bar chart to navigate to support and disagreement category comments. They were interested in seeing the polarizing views that the paper generated. One participant said they would like to navigate to more debates via comment threads. However, the sentiment axis was not commonly used, even though it also served to find positive and negative comments. Participants explained they did not trust the calculations performed by the sentiment axis. These results indicate that human-selected categories might be more trusted than AI generated scores.

4 DISCUSSION

With AnnoDiver, we observed that adding visualizations into a social annotation system did seem to affect the engagement with comment exploration and creation. For *HI*, our quantitative results indicated some promise in that the experimental group tended to view more annotations. Qualitative feedback indicated that this might have been due to the bar chart helping users navigate between polarized views via support and disagreement categories. Although usage of our prototype in the 8-person user study did not show support for H2, the visualizations seemed to have led to an increase in the creation of comments. Further work could explore whether these trends hold in a larger study run in a more natural setting with a wider range of participants, and if so, what specific attributes of the system might be responsible.

ACKNOWLEDGMENTS

We would like to thank Dr. Maneesh Agrawala and the teaching assistants from CS 347 at Stanford University who provided feedback on our project, as well as our participants.

REFERENCES

- [1] Siti Salwah Salim. 2015. A systematic review of shared visualisation to achieve common ground. *Journal of Visual Languages & Computing* 28 (2015), 83–99. Publisher: Elsevier.
- [2] Julie A Murphy. 2021. Collaborative Annotation: Tools for Enhancing Learning and Scholarly Communication. *Serials Review* (2021), 1–6.
- [3] Fei Gao. 2013. A case study of using a social annotation tool to support collaboratively learning. *The Internet and Higher Education* 17 (2013), 76–83.